Appendix 1: Model-based Pn and Sn path-dependent travel-time corrections for IMS stations in the Mediterranean, North Africa, Middle East, and Western Eurasia

This Appendix is to support the recommendations in this proposal that the regional path-dependent Pn and Sn travel-time corrections for IMS stations in the study region be incorporated into the operational location process. In this Appendix we summarize the 3D model, SSSCs, and modeling errors developed, and the scientific justifications based on validation testing using JHD empirical path corrections and event relocations. More details on each of the subjects are given in Appendices 2-9.

Introduction

The Group-2 Consortium carries out seismic location calibrations for IMS stations in the Mediterranean, North Africa, Middle East, and Western Eurasia using 3D models. Source Specific Station Corrections (SSSCs) for IMS seismic stations are developed to improve location accuracy and reduce error ellipses. We define the study region as a rectangle covered by the 20 degree circles around 32 designated Group-2 stations, -20° to 80° North and -40° to 100° East (http://g2calibration.cmr.gov/calibration/result.html). Our goal is to develop SSSCs for Pn, Sn, and Lg phases out to 20° and the Pg phase out to 8° for all IMS stations in the study region in two phases of the work. In Phase 1 we develop Pn and Sn SSSCs for a source depth of 10 km (McLaughlin, 2000). The current IDC software cannot handle depth-dependent SSSCs, and it applies SSSCs at all source depths. The 10 km source depth is a compromise for all regions of crustal seismicity. In Phase 2 we will refine and improve the models and methods to obtain SSSCs, including depth dependence and Pg and Lg SSSCs. In both phases, the model-based corrections are compared with empirical path corrections, and GT events are relocated to demonstrate improvement in event locations when using the corrections. This document reflects the Phase 1 results on data, model, and validation testing, and improvements and refinements are expected in Phase 2.

During Phase 1 a 3-D velocity model, CUB1.0, was developed and SSSCs were computed and tested. The CUB1.0 is constructed by combining 3D global mantle models with global crust and upper mantle models using improved group and phase velocity data sets and inversion methodology (Shapiro and Ritzwoller, 2001; Ritzwoller et al., 2001) (Appendix 2). SSSCs were computed using ray tracing (Barmin, 2001) (Appendix 3) for IMS stations, IMS surrogates, and other non-IMS stations for validation and testing. Regional SSSCs are defined on rectangular 1° x 1° latitude/longitude grids where both a travel-time correction and a modeling error are given at each grid point. The corrections are relative to the IASPEI91 travel times used in the PIDC/IDC location process. Modeling errors, the uncertainty in the predicted travel times, are estimated to ensure 90% ellipse coverage. In Phase 1 modeling errors are estimated from the variance of the travel time residuals (Bhattacharyya et al., 2001) (Appendix 4), with respect to the 3D model, in the EHB catalog (Engdahl et al., 1998). Comparisons were made between the model based path corrections and empirical path corrections for cross-validation and model error evaluation (Israelsson et al., 2001a) (Appendix 7).

GT0-GT10 events in the study region have been collected for relocation testing (e.g. Bondár et al., 2001) (Appendix 6). Validation testing of the model-based SSSCs is conducted by relocating mostly GT0-GT5 events. GT10 events may also be used to extend the coverage of the region where GT0-GT5 events are not available. Events were relocated with and without Pn and Sn SSSCs and results were compared (Yang et al., 2001a; Ritzwoller and Levshin, 2001) (Appendices 8-9). Detailed evaluations of location improvement/deterioration were conducted using a suite of evaluation criteria and metrics. Online testing was also successfully conducted for the CUB1.0 SSSCs (Yang and McLaughlin, 2001; Oancea and Caron, 2001) (Appendix 5).

Pn and Sn SSSCs from Ray tracing the 3D CUB1.0 Model

Continent- and even global-scale 3-D models of the crust and upper mantle have been developed recently with sufficient reliability and resolution to improve event locations using regional phase data. A global 3D S velocity model ($2° \times 2°$), CUB Model 1.0, was constructed by simultaneously inverting broad-band group and phase velocities (Shapiro and Ritzwoller, 2001) (Appendix 2). Monte-Carlo inversion of surface wave dispersion curves was performed to obtain the shear velocity model in the crust and upper mantle, together with uncertainties. Constrained by a priori information, the inversion yields an ensemble of acceptable models at each spatial node in a Monte-Carlo sampling of model space. The estimated model is defined by the median and the uncertainty estimates is given by the half-width of the ensemble. Robust features of the model (<200-250 km) are identified using the uncertainty estimates. This method is applied to a large data set of fundamental mode surface wave group and phase velocities, including Rayleigh group velocity (18-175 s), Love group velocity (20-150 s), and Rayleigh and Love phase velocity (40-150 s). Over 100,000 group velocity paths and 50,000 phase velocity paths were used in the inversion. Data coverage is better for Rayleigh waves, intermediate periods, and in the northern atmosphere. The coverage is the best in Eurasia and the worst in Africa, central Pacific, part of the Indian Ocean, and Antarctica.

The starting model of crust and upper mantle was built upon several global models and regional information for Eurasia. From the UCSD sediment model (Laske and Masters, 1997) and CRUST5.1 crustal model (Mooney et al., 1998), the initial CUB crustal model includes a water layer where appropriate, topography in the solid surface and Moho, and 3D P and S velocity variations in the sediments and crystalline crust. The initial upper mantle shear velocity model was based on the S20A model (Ekström and Dziewonski, 1998) modified with radial anisotropy from PREM (Dziewonski and Anderson, 1981), with an average velocity from the AK135 model (Kennett et al., 1995) to remove the 220 discontinuity in PREM. The upper mantle velocities of the CUB Model were constrained by Pn maps (Rizwoller el al., 2001). From the CUB shear velocity model the P velocities were calculated:

$$\frac{Vp - Vpak135}{Vpak135} = 0.5\frac{Vs - Vsak135}{Vsak135}$$

where Vp and Vpak135 (Vs and Vsak135) are P (S) velocities from the CUB and AK135 models, respectively. The CUB 1.0 Model is radially anisotropic in the uppermost mantle (~200 km) and isotropic elsewhere. The resolution of this model is 400 km laterally, and 5 km in the crust and 100 km in the mantle vertically. Figure A-1 shows sample slices of the 3D model.

SSSCs are travel time corrections of the 3D model relative to the IASPEI91. They are defined on a latitude/longitude grid surface for a given phase, depth, and station. In Phase 1 we developed regional Pn and Sn SSSCs from the CUB 1.0 Model for all IMS stations in the Group-2 study region out to $20°$. The SSSCs are for fixed depth of 10 km for all stations. A 2D ray tracer was used for numerical calculation of travel times for refracted and reflected P or S waves in 3D laterally inhomogeneous media along 2D cross sections of a spherical earth (Barmin, 2001) (Appendix 3).

The input data of the ray tracer include the 3D model, station information, and desired azimuth and depth. The 3D model is represented by 3D grids in geographic spherical coordinates with equal lateral cell (latitude and longitude) and variable depth. At any given geographic point the model is defined by a 1D velocity profile consisting of a set of depth cells with a constant velocity at each depth cell. The CUB model is given on a grid cell spacing of $2°$ and 8 depth layers:

- Layers 1-2: Soft and hard sediments
- Layers 3-5: Upper, middle, and lower crust
- Layer 6: Upper mantle from Moho to 400 km discontinuity
- Layer 7: Transition zone (400-670 km)
- Layer 8: Lower mantle (670-2740 km)

A 2D cross section is constructed from the 3D model along the given azimuth and earth's flattening for the 2D model is performed using a logarithmic conform transformation. The ray field is computed by solving the Eikonal equation using the Runge-Kutta 4th order method with step doubling. It includes direct, diving, and head waves defined by groups to determine the behavior of each ray at the incidence point. The ray tracer computes the ray and time fields of diving waves for separate layers, and joint travel time tables for the groups. By interpolating the predicted travel times, SSSCs relative to IASPEI91 are then computed at any given depth between 0 and 200 km out to $20°$ from the station for a given phase. The known deficiencies of the ray tracer include shadow zones, sometimes unstable SSSC behavior in some areas where model parameters vary strongly, and grouping of layers required for computing SSSCs. Figure A-2 shows an example of ray tracing calculations and Figure A-3 shows an example of SSSCs.

We compared the SSSCs at a source depth of 10 km with surface-focused SSSCs. The difference between the two is usually no larger than a few tenth of a second which is well below the estimated uncertainties. For 18 randomly selected IMS stations in the Group-2 study region, the median difference is 0.2 sec with a median absolute standard deviation, normalized to a Gaussian distribution, of 0.25 sec.

Modeling Errors of the Pn and Sn SSSCs

As for SSSCs, modeling errors associated with the SSSCs are also defined on the latitude/longitude grid surface for the give phase, depth, and station. These error estimates are important because they are used, together with measurement errors, to weigh phase picks in event locations and to construct error ellipses. Ideally the model errors should reflect the uncertainties in the 3D model, so they should be azimuth- and station-dependent. Currently research is being carried out

on how to formulate, propagate, and utilize model errors in model development and event location.

In Phase 1 we developed empirical model error surfaces (Bhattacharyya et al., 2001) (Appendix 4) using travel time misfits obtained from the EHB catalog (Engdahl et al., 1998). Event were selected from the EHB data set that are considered to be accurate to within 15 km (i.e., GT15 events). Ray tracing was performed to predict the travel times of over 1,000,000 Pn rays. Misfits were obtained by comparing the predicted travel times with the observations from the EHB bulletins as the AK135 and IASPEI91 travel times are the same for regional phases. Model error for Pn were estimated from the standard deviations of the Pn misfits as a function of distance. Sn error estimates were obtained by scaling Pn errors by a factor of two. In general, [δVs/Vs] ~ 2.0 [δVp/Vp] (Robertson and Woodhouse, 1995; extrapolated to upper mantle bottoming depths) and we expect the errors to scale similarly. Figure A-4 shows average L1-norm misfits as a function of the epicentral distance for all of the stations across Eurasia and for IMS stations and surrogates only. The average misfit grows nearly linearly until about 15˚ and then decreases. This 10˚ - 15˚ feature results predominantly from errors in predicting the distance where Pn transitions to a diving P. This transition is seen in the SSSCs as a high velocity ring that occurs between 10˚ and 15˚ at most azimuths. Predictions of the onset of the Pn to diving P transition are very sensitive to the upper mantle vertical gradient. Therefore, small changes in the model produce moderate changes in travel time predictions in the 10˚ - 15˚ range.

A series of analyses support the distance-dependent model errors. First of all, we compared these estimates with the modeling error baseline used in the IDC event location process, which is also azimuthally invariant and distance-dependent only. Although the methodologies in deriving the two sets of error estimates are similar, the IDC errors were derived using a background 1-D IASPEI91 model while the new errors are based on the 3D CUB1.0 model. With better estimates of the lateral variation of structure, we expect that the level of un-modeled signal in our estimates to be smaller than those obtained using IASPEI91. As shown in Figure A-5, our model errors are significantly smaller than the IDC's, and the shapes of both error curves are similar. Secondly, to limit the effect of errors in the source location and origin time on our analysis, we also estimated the model errors, relative to the IASPEI91, using only EHB GT5 events. As expected, the EHB GT5 error estimates are similar to the IDC baseline for Pn (Figure A-5). The Sn errors differ slightly in amplitude and phase with a prominent peak value at 16˚ for the EHB GT5 events. This result independently confirms the general amplitude and distance dependence of modeling errors in the current practice at IDC.

We also checked the internal consistency of the EHB catalog used in our error estimates by comparison with Pn and Sn time residuals from the PIDC/IDC REB. Stations are limited within the Group-2 region and events are selected with at least 15 defining phases and depth less than 40 km (to limit P/Pn identification ambiguity). We compared the standard deviation of the travel time residuals as a function of distance. For Sn it is similar to those for the EHB GT5 and IDC modeling errors. The Pn errors are significantly smaller than the corresponding ones for EHB GT5 and IDC, probably due to the analysts' intention for better fit with IASPEI91 in locating events. This experiment suggests that the observed travel time can be a function of the underlying velocity model. Therefore, using a more realistic 3-D model, especially in the analysis of regional arrivals, can improve phase picks and event location.

One of the components of the modeling error is the structure that has not been accounted for in the 3-D models and can bias travel time predictions from SSSCs computed for those models. Some indicators of this unmodeled signal can be obtained from comparing different 3D velocity models. We compared the travel time predictions for identical paths for the CUB1.0 model and another model developed in Phase 1, the SAIC-HRV model. The latter is a regionalized crustal model, developed from published tectonic maps and 1D velocity models, combined with a 3D mantle model parameterized in terms of radial and horizontal cubic splines using a combination of direct and differential travel times and surface-wave phase measurements (Antolik, 2001; Bondár, 2001); this model is being refined in Phase 2. We compared the corresponding Pn and Sn SSSCs for a given station for more than 750 stations. The test is limited since we use only two models, and both are optimally smooth and thus error estimates from this test are conservative. Comparisons show that velocities in the SAIC-HRV model are generally faster (Figure A-5). The variation of the travel time differences between the two models as a function of epicentral distance shows that the Pn values are lower than the Phase 1 modeling errors. This comparison supports the hypothesis that the former explains a fraction of the total error budget as the latter also contains the effects of measurement error, origin time errors and the effect of smaller scale structure not modeled by either of these two models.

Validation Testing Using Cluster Analysis

Empirical station path corrections were obtained from event cluster analysis using Joint Hypercenter Determination (JHD). The event cluster analysis serves as a validation tool of three areas of the location calibration: (1) GT events, (2) 3D velocity models by comparing empirical and model based corrections, and (3) model errors. The empirical corrections were calculated relative the IASPEI91 standard travel time tables for first arrival P phases at distances between 2° and 20° using the procedures described by Israelsson (2001a). Depths of all events were held fixed at that of the GT event during the JHD processing. The JHD processing was validated with a few clusters which included several GT0-GT1 events (Israelsson, 2001a) and with an independent cross-validation with the Hypocenter Decomposition (HDC) cluster analysis (Israelsson et al., 2001b).

The focus of the compilation of event clusters is the European part of the study area using the Group 2 GT0-GT5 events. Attempts were also made to generate additional GT events from the resulting JHD solutions (Israelsson, 2001b; Israelsson and Hofstetter, 2001). In addition, event clusters were defined without available GT information in order to extend the geographical coverage in the Mediterranean region (assigned GT25). The compilation and JHD processing resulted in 47 event clusters, 19 with GT0-GT5 and 28 with GT25 (Table A-1; Figure A-6). As a result of the JHD the diameter of more than half of the clusters shrank and the scatter of the epicentres (tightness) was reduced for about 75% of the clusters. The coverage is dense across the Mediterranean region and more sparse in northern and eastern Europe.

Figure A-7 shows the distribution of all JHD path corrections. The corrections were normalized for each cluster by subtracting the median of the path corrections for stations at distances between 30° and 90°, if there are teleseismic stations. Most JHD path corrections are within 5 sec and approximately Gaussian. The median of the corrections is slightly positive (0.6 sec), indicating slower travel times than that of IASPEI91. The overall spread in the path corrections is

characterized by a standard deviation of 1.53 sec. This is clearly a smaller scatter than one would expect from the model errors used for the IASPEI91 tables. The normalization applied to the corrections described above could have narrowed the scatter as the JHD determination as a boundary condition set the average of corrections estimated to each cluster to zero. Fast paths are concentrated in the northern shield/platform regions and slow paths in the southern tectonic regions. Figure A-8 shows the corrections as a function of distance. The corrections show similar ranges except for distances around 15° where the range becomes larger. The variation in standard deviation is similar to that of the standard deviation assumed for the CUB1.0 model errors.

Event clusters located close to one another allow comparisons of consistency in the estimated path corrections. Comparisons of the JHD path corrections of some close clusters across the study region show general agreement. Overall there is good correlation for separations less than 1° between the cluster centers with the same station. A comparison of the correlation between JHD path corrections can also be made for station pairs as a function of station separation for the same cluster. As expected, there are similar features as for the pair-wise cluster comparison. The high correlations exhibited for the empirical path corrections both among clusters and among stations for distances less than 1° lend support to use of event clusters without GT information (GT25 events in Table A-1) and to a 1° sampling for SSSCs. The correlation distance for the empirical corrections suggest that a resolution of 2° or 4° for a velocity model might result in under-sampled SSSCs.

We compared the JHD corrections with SSSCs calculated from the CUB1.0 model (Shapiro and Ritzwoller, 2001; Ritzwoller et al., 2001). The comparison is limited to first arrival P between 2° and 18° (in all 3890 paths). The CUB1.0 SSSCs (zero depth) were calculated for each cluster centered at the location of the event held fixed in the JHD location (zero elevation). Some differences have been noted in SSSCs as a function of shallow depth, so this simplification is bound to introduce some scatter in the data. Furthermore, the SSSC values used in the comparisons were extracted with limited accuracy without interpolation from a polar co-ordinate output format with distances to the nearest 25 km and azimuths to the nearest 3 degrees. Possible differences in ellipticity and elevation corrections were not accounted for in the comparison, but they were small for the paths analysed (ranges of -0.37 - 0.26 and -0.04 - 0.46 for ellipticity and elevations, respectively). Another limitation in the accuracy of the comparison is imposed by the uncertainty of the origin time of many of the events held fixed in the event clusters. Such origin times determine the level of the JHD path corrections, so a shift in the origin time introduces a corresponding shift in the path corrections. Hence, strict comparisons can only be made in a relative sense for each cluster. In comparisons based on data of more than one cluster the corrections were normalized; each set of JHD corrections and SSSCs for a given cluster was normalized to a zero median value. Figure A-9 shows an example for a JHD cluster in Mascara, Algeria. There is good correlation between the empirical and model based path corrections. This figure shows another feature typical of most clusters, the JHD corrections span a larger range than the corresponding CUB1.0 values (slope < 1). In this figure the JHD corrections are also plotted on top of the CUB1.0 SSSC map, and the high correlation between the two types of correction result in general overlap of the colors.

Table A-2 lists some comparison statistics for each cluster including median length of paths in degrees, range and standard deviation of JHD and of CUB1.0 values, correlation coefficient, and

the standard deviation of the differences CUB1.0-JHD. The JHD corrections and the CUB1.0 for all the clusters combined yield a correlation coefficient of around 0.3. The data in Table A-2 shows that there is a fair (> 0.3) or high correlation for about half of the clusters (21 out of 43). Pearson's correlation test was applied to each cluster. The null hypothesis assumes that the two types of corrections come from un-correlated data while the alternative hypothesis assumes positive correlation. The low p-values suggest correlation in most cases; only for 6 clusters is the p-value larger than 0.5. Some of the low correlation values appear to be an effect of the distance distribution of the paths which are dominated by paths lengths around 5°. For shorter paths correlations appear lower, where the correlation coefficient is shown as a function of distance. The data were separated into two sets, one included only clusters for which the correlation was higher than 0.3 and the remaining clusters made up the other set. For both sets there is a clear change in the correlation coefficient with low values at distances less than about 5°. Apart from shorter distances the empirical JHD corrections and model CUB1.0 values show fairly consistent correlation. There are also exceptions from the distance effect. For example, the event cluster at Annecy, France, dominated by near stations, shows high correlation, even if the ranges of the JHD corrections are much larger than those of the CUB1.0 values.

There is a geographical variation of the correlation (distance effect "removed"). Agreement between JHD and CUB1.0 is highest for clusters in the northern Europe. These are also mostly of GT1 category and surface events; for four out of six clusters were at zero depth. It is unlikely that the GT category should affect the correlation as JHD corrections correlate up to 50 km or more. The depth, however, could be a contributing factor to lower correlations for the shallow earthquake clusters. All clusters with low correlations are in Greece, Italy and Northern Africa.

Standard deviations of the differences in CUB1.0 SSSC - JHD path corrections are listed in Table A-2 for each cluster. The values range from 0.56 to 1.92 with a median of 1.15. Figure A-10 shows the bulk distribution of all the differences in JHD-SSSC which is approximately Gaussian. The median for the bulk distribution of 1.15 is in agreement with the median among the clusters. This agreement supports the normalization used in the comparison of JHD and CUB1.0 residuals, as the median for the clusters is independent of the normalization while the median for the bulk is not independent. Clusters with high correlations also generally have low standard deviations. The median of standard deviations can be compared with the standard deviation of JHD values, 1.53 (see Figure A-7), which represents a measure of the fit of the IASPEI91 tables to empirical travel times. The reduction of this measure by CUB1.0 from about 1.53 to 1.15 suggests improvement in calculated travel times with the CUB1.0 model relative to empirically observed times. This corresponds to and overall variance reduction of 44%, assuming that the normalization of the JHD corrections is unbiased and may be an overestimate as the median of the variance reduction of the individual clusters is about 20%.

The standard deviations of the CUB1.0 SSSCs (see Figure A-8) and the estimated uncertainties for the JHD corrections can be accounted for in the comparisons between the JHD corrections and SSSCs. The 90% confidence intervals overlap in more than 99% of the cases. Thus, accounting for uncertainties in both JHD path corrections and SSSCs results in good agreement between the two types of correction, although correlation values in many cases are low. Indeed, it appears that the CUB1.0 model errors may be on the conservative side as indicated above. If we use confidence intervals with the CUB1.0 distance dependent model errors that are reduced by a

factor of two there is overlap in about 80% of the cases, a degree of overlap one would expect for two independent 90% confidence intervals. The lower diagram of Figure A-8 shows the standard deviation of the differences CUB1.0-JHD as a function of distance and the CUB1.0 model error curve is again included for comparison. Apart from distances around 15° the CUB1.0 error is higher than the standard deviation of the differences.

Event Relocation Validation Testing and Test Data Sets

To verify event location improvement relative to the GT and location error ellipse coverage, validation testing is conducted by relocating events using SSSCs in the Group-2 region. Events used in the testing are GT0-GT10 events and are not directly used in the model development. SSSCs from a large set of stations are applied in event location to validate the models and model errors. To assess IMS location improvement SSSCs are applied to only IMS stations and surrogate stations. The effect of mixing calibrated and uncalibrated regional and teleseismic data is also tested. IMS surrogate stations are used to simulate the IMS network where future IMS stations are not yet deployed and/or data are not available from existing IMS stations for testing. They are existing stations within 75 km of the corresponding IMS stations.

As shown in Table 2 of the main text, four data sets are used in event relocation tests, including (1) the Fennoscandian GT events used in a previous study (Yang and Mclaughlin, 1999), (2) high-quality Group-2 GT0-GT10 events, (3) GT10 Mid-Ocean Ridge and Transform (MORT) events, and (4) GSETT-3 REB events. Since the expected improvement in location is on the order of 10 km, the location of reference events should preferably be within 5 km accuracy or better. To extend coverage of the region, however, other events may also be used where GT0-GT5 events are not available. We conduct thorough testing using the high-quality GT0-GT10 events (Data Set 2) for relocation and error ellipse validation. There are about 600 such events, including nuclear explosions, chemical explosions, and well-located earthquakes, particularly from JHD/HDC cluster analyses (Engdahl and Bergman, 2001). Events with at least three Pn (Sn) arrival data are used in validation testing. 96% of the events are GT0-GT5 events and 85% of the events have depths less than 10 km. Figure A-11 shows the event distributions of Data Set 2. The Fennoscandia data set (Data Set 1) was previously used in testing 1D SSSCs for the Fennoscandian stations. It mainly contains 425 GT events in Fennoscandia, originally considered as GT2. Among them 181 events are in the Group-2 GT0-GT10 data set and 50 events are now classified as GT5. The MORT data set (Data Set 3) include 35 events in the Gulf of Aden and North Atlantic. These events are referenced to the bathymetric features and are estimated as GT10. Only a subset of the events can be located due to phase association problems between P and Pn. In the GSETT-3 REB event data set (Data Set 4) over 6000 events are used for testing to assess the real world situation when SSSCs are applied.

We evaluate statistics on mislocation, error ellipse area, 90% error ellipse coverage, origin time differences from GT, origin time error, and standard deviation of observations. Our major evaluation metrics include those recommended by the 1999 Oslo Location Workshop (CTBT/WGB/TL-2/18, 1999). Additional metrics are also developed to measure the performance of the SSSCs. Both the L1 norm (median, spread, minimum, maximum, 20th, 40th, 60th, and 80th percentiles) and the L2 norm (mean, variance, standard deviation, and average deviation) are calculated for the distance from GT, size of error ellipse, 90% ellipse coverage, origin time, origin time error, and

misfit obtained with and without SSSCs. Student and Wilcoxon significance tests of paired samples are also applied, and statistics significant at least at 95% level are included in this document. Besides applying the evaluation metrics to an entire data set, we further divide the relocated events into several classes based on the GT accuracy (i.e. within vs. beyond the GT accuracy when located with and without SSSCs), and on mislocation (i.e. within vs. beyond 18 km). We also compare the numbers of events in a data set, with and without SSSCs, which satisfy the "trinity criteria", i.e. located within 1000 km^2 error ellipses that contain the GT and within 25 km distance from the GT.

Event Relocation Test Results

As shown in Table 2 in the main text, ten relocation tests were conducted using four test data sets.

Benchmark Test Using Fennoscandian GT Events (Data Set 1; Test 6)

The Fennoscandian data set and 1D SSSCs that are currently used in the PIDC/IDC system serve as a benchmark of the SSSCs developed using the 3D model by ray tracing. The location results for the GT events are similar when the CUB and 1D Pn and Sn SSSCs are applied. With the CUB SSSCs, the median mislocation is improved for 60% of events by a median 10 km, and the median ellipse area is reduced by 1900 km^2 (from 3700 to 1800 km^2) without loss of the 90% coverage (99%). With 1D SSSCs, the mislocation improvement is similar, and the median error ellipse area is reduced to 1100 km^2 with 94% coverage, compared to 98% coverage without SSSCs. The differences in error ellipses and 90% coverages between the two set of SSSCs resulted from the differences in modeling errors. The Pn modeling errors for the CUB model are about 1 sec larger than the 1D's within 12° and about 0.5-1.0 sec smaller beyond 12°. In the testing about 80% of the data are Pn defining phases.

Direct comparisons were also made between the 3-D and 1-D SSSCs for Fennoscandian stations. Differences between the two as a function of distance are normalized by the distance-dependent modeling errors, σ. The percentages of the difference below 1σ and 2σ levels indicate how well the difference can be accounted for by the modeling errors. Figure A-12 shows the histograms and cumulative distributions for both Pn and Sn differences. The average misfits of the differences between the two sets of SSSCs are -0.25 sec for Pn and -3.61 sec for Sn, with the CUB model being faster. 84% and 100% of the Pn differences are below the 1σ and 2σ levels, respectively (48% and 93% for Sn). The two sets of Pn SSSCs are similar, with 84% of the differences within the CUB modeling errors. While the Sn differences are significantly larger than the Pn's, the Fennoscandian Sn SSSCs are four times larger than the Pn SSSCs (29 vs. 7 sec). The CUB velocity model is a robust S model scaled to P assuming velocity variations between the CUB model and AK135 have a factor of 2 between P and S. The 1D S velocity model is simply scaled using the depth-dependent IASPEI91 S to P ratio from the 1D P velocity model, which was developed based on Pn observations from historical Deep Seismic Sounding Profiles (Bondár and Ryaboy, 1997).

The relocation results show that the new 3D SSSCs perform similarly to the 1D SSSCs in Fennoscandia. Replacement of these 1D SSSCs with the consistent set of 3D SSSCs for all of Europe would do no harm to the system and all SSSCs would be based in a consistent 3D model.

Model Validation Testing Using GT0-GT10 Events (Data Set 2; Test 2)

Model validation testing is conducted using this data set of high-quality GT0-GT10 events. Depth is fixed to zero in all relocation tests since these events are mostly shallow. We include all locatable events in the evaluation. Events near the boundary of the study region may be poorly located since we do not use any stations outside the study region. A total of 571 GT0-GT10 events are relocated using only Pn and Sn phases, with and without SSSCs, for all stations. As shown in Table A-3 the relocation results are similar between all GT0-GT10 events and GT0-GT5 events since there are only 24 GT10 events (all in the Aden and Koyna clusters). The median improvement in mislocation is larger for the GT0-GT2 events (10 km) while smaller for the GT5 events (6 km). Figures A-13 to A-17 show maps of event relocations.

Relocation results of the 571 GT0-GT10 events shows that (Table A-3), with SSSCs,

- 60% of the events are improved by a median 7.9 km, and 47% of the events are improved by more than 20%. 40% events are deteriorated by a median 6.4 km, and 31% deteriorated by more than 20%.
- Overall median mislocation improved by only 14% (16.5 to 14.1 km), but is statistically significant at 95% level. The 80th percentile mislocation decreased 33% (from 43 km to 29 km).
- 63% of 57 GT0 events were improved. 58% of the GT0 improved by more than 20% compared to only 28% of the GT0 events deteriorated by more than 20%.
- 46% more events (41 vs. 28) were located within GT accuracy. 34 events moved from outside to inside GTX accuracy compared to 21 events moved from inside to outside GT accuracy. 14% more events (361 vs 317) located within 18 km.
- 100% of events have reduced error ellipses. Median ellipse area is reduced 51%, from 4600 to 2240 km^2.
- 90% ellipse coverage reduced from 97% without SSSCs to 91% with SSSCs.
- Coverage was closer to the theoretical χ^2 distribution for the 20th, 50th, 80th and 90th percentiles.
- 74 more events satisfied the “Trinity criteria” (increase from 11% to 24%) (error ellipse area < 1000 km^2, covers GT, mislocation < 25 km).
- Origin time errors were reduced for 99% of events. The median improvement, relative to the "GT" origin time, is 0.9 seconds (29% improvement), from 3.1 to 2.2 sec.
- Standard error of observations improved for 61% events with 17% median improvement.

Figure 4 in the main text shows mislocation improvement scaled by GT accuracy where GT0 events are assigned a 1 km uncertainty. About 31% of the events moved only within GT accuracy or less than 1 km. 40% more events are improved than deteriorated. Figure 5 in the main text com-

pares mislocations of the GT0-GT10 events relocated with and without SSSCs. Overall mislocations are reduced with SSSCs, particularly in the tails of the distribution. An analysis was performed (McLaughlin and Bondár, 2001) on the expected degradation versus improvement at selected test coverage values conditioned on the performance of the reference event data set and the total model error (GTX + measurement + model). At every level of test coverage between 0.1 to 2.0, the expected number of events that got significantly worse is never more than should be expected by random chance.

Figure 6 in the main text shows comparisons of the coverage parameter (mislocation normalized to the error ellipse), with and without SSSCs, by GTX category, compared to the χ^2 distribution. The coverage statistic is χ^2 distribution with an expected 90th percentile of 1. It measures how well the 90% error ellipse covers the GT location. There are large deviations from the expected values, particularly for the uncalibrated case. The single point comparison of the 90% coverage parameter is above 90% both with and without SSSCs. Coverage is generally better than expected below the 90th percentile and worse above the 90th percentile. This indicates that *a priori* errors are over estimated for the majority of events but under estimated for a small number of events. As shown in this figure, median coverage was 0.1, significantly less than the expected value of 0.3 (off by 67%). This indicates that the majority of events were actually located closer to GT than should be expected from the model and measurement errors. The 90 percentile coverage was 0.9, close to the expected value of 1.0 (off by 10%), demonstrating that the 90% error ellipses were indeed "honest". However, the number of outliers (18 events with coverage larger than 2) clearly exceed the expected number at a high significance level (5 events). This indicates that the underlying "Gaussian statistics" for model and measurement error are probably inadequate for this data set. The modeling errors appear to be a conservative compromise under the condition that 90% error ellipses are "honest". However, in order to predict "honest" 95% or 98% error ellipses the modeling errors would need to be inflated. Given the error model, 50%-60% of the time the location procedure performs better than should be expected.

There are large error ellipse reductions for events with low ndef, number of defining phases (Figure A-18). Such events tend to be poorly located, both with and without SSSCs (Figure A-19) with similar 90% coverages (Figure A-20). Event relocations are within 90% coverage as long as the values are no more than 1 in Figure A-20. As shown in Figure A-19, mislocation improvements are generally larger than deteriorations up to azimuthal gap (azgap) of 230 degrees. Very often events with small ndef tend to have large azgap, but events with large ndef may also have large azgap (Figure A-21). For events with low ndef/large azgap, while mislocations may not necessarily be improved using SSSCs, error ellipse areas are significantly reduced without degradation of event location and 90% coverage.

The test results of 571 GT0-GT10 events using Pn and Sn phases show that, with SSSCs, more events are improved than deteriorated and improvements were larger than deteriorations. The mislocation reduction is statistically significant, and the degradation is less than expected by random chance more than 90% of the time. The SSSCs significantly reduced error ellipses without loss of 90% coverage. The metrics met or exceeded the evaluation criteria of the 1999 Oslo recommendations (CTBT/WGB/TL-2/18, 1999).

An independent relocation testing was conducted by the University of Colorado at Boulder using an L1 norm grid search method. A total of 150 GT5-GT10 events from 15 event clusters obtained from cluster analysis (a subset of Data Set 2) were relocated using Pn phases with the CUB 1.0 Model and with the 1-D AK135 model, respectively for comparisons. Test results show significant improvement in event location using SSSCs as a function of ndef. When ndef>15, mislocations are improved for 75% of the events, from 19 km using AK135 to 11 km using CUB1.0. The quality of locations depends strongly on the data rejection criterion.

Model Validation Testing Using MORT GT10 Events and GSETT-3 REB Events (Data Sets 3 & 4; Tests 4 &5)

We test using MORT GT10 events to extend path coverage into the ocean basins since no GT0-GT5 events are available there (see Figure A-11). A total of 24 MORT GT10 events are relocated with and without Pn and Sn SSSCs, but the location algorithms are very sensitive to outliers in the distance range of 15°-20° due to misassociation of P to Pn. Therefore we relocated MORT events using arrivals within 15° and 15 events were relocated. The majority of events have significant reductions in both mislocation and error ellipses using SSSCs. However, there are large mislocations and poor ellipse coverage, possibly due to poor station geometry.

For the 15 MORT GT10 events relocated using arrivals within 15° with SSSCs,

- The median mislocation is reduced from 98.6 km to 82.2 km.
- 80% events are improved by a median 21.5 km, and 20% events are deteriorated by a median 10.8 km. 40% events are improved by 20% or more, compared to 7% deteriorated by 20% or more.
- Median error ellipse areas reduced from 22000 to 12430 km^2.
- 90% coverages are reduced from 87% to 73%.
- The median 90% coverage parameters with and without SSSCs are similar to the expected values, but larger beyond the 50th percentile.

Figure A-18 shows the locations of GT10 MORT events along the Mid-Ocean Ridge relocated using Pn and Sn phases from all stations within 15°, with and without SSSCs. Most events are closer to the Mid-Ocean Ridge with SSSCs than without calibration.

To test the SSSCs in the real world situation, we used 6835 GSETT-3 REB events between 01/01/1995 and 02/20/2000 with shallow depth (< 40 km) and at least 3 Pn (Sn) phases in the region of -10° to 80° N and –30° to 70° E. Events were relocated using Pn and Sn phases, with and without SSSCs. Different from other offline relocation tests, in this testing slowness and azimuths were also used, with SASCs (Slowness Azimuth Station Corrections), if they were defining in the REB. About 1/3 of the defining phases in the data set are azimuth defining (40% of these with regional SASCs) but almost none are slowness defining. A statistical test was conducted to test the hypothesis that seismicity would be more clustered with calibrations. 899 events were located only with SSSCs only, and 270 events only without SSSCs. For the 4786 GSETT-3 REB events located both with and without SSSCs,

- The location differences are small, but significantly more events were located with SSSCs indicating that the better travel time and error predictions facilitates more consistent location estimates. Distances between with and without SSSC locations: median=0.08 spread=0.05 smad=0.07 min=0.00 max=5.09, where spread is the median of absolute deviation, and smad is the median of absolute deviation normalized to a Gaussian distribution.
- The median nearest neighbor distance (a measure of cluster tightness) was decreased. For events within 5 km of other events (r<5), the nearest neighbor distance decrease was statistically significant. Nearest neighbor distance without SSSCs for r<5: median=1.53 spread=0.88 smad=1.31 min=0.00 max=5.00. Nearest neighbor distance with SSSCs for r<5: median=1.47 spread=0.87 smad=1.29 min=0.00 max=4.99.

IMS Location Improvement Testing (Data Set 2; Tests 7, 8, 9, 10 &11)

IMS location improvement is simulated using the Group-2 GT0-GT10 events with IMS stations and IMS surrogate stations. To further assess the effect of SSSC calibrations in operational event location, we tested for both calibrated phases only, and mixing calibrated with uncalibrated phases, including both regional phases and teleseismic phases. In each case events are relocated using IMS stations only and IMS + surrogate stations.

Calibrated Pn and Sn phases only

A total of 240 GT0-GT10 events were relocated using only IMS stations, with and without SSSCs:

- 62% of events improved by a median 7.6 km compared to 38% deteriorated by a median 8.0 km.
- 49% of the events are improved by 20% or more compared to 31% deteriorated by 20% or more.
- 90% error ellipse coverage reduced from 100% to 98%.

The statistics are similar to the previous cases where all stations are used, but the median deterioration is slightly larger. Some events are poorly located because fewer defining phases are applied, azimuthal gaps are large, and station geometry is poor. The mislocation distribution for this test set of 240 events limited to the IMS stations is not distinguishable from the larger set of 571 events using all stations. For example, 20% of the events are mislocated more than 41 km without calibration compared to 32 km with calibration (22% reduction). These numbers are not statistically significantly different from the 80th percentiles of the larger data set. Again, the most important improvements occur in the tails of the mislocation distribution. When we consider the non-parametric statistics; 38% (11 vs. 8) more events are located within GTX accuracy with calibration, 19% (150 vs. 126) more events are located within 18 km with calibration. Calibration does more improvement and “does no harm”. These results demonstrate SSSCs should improve locations based on regional data for a limited IMS network.

A total of 318 GT0-GT10 events were relocated using only IMS + surrogate stations, with and without SSSCs:

- 59% of events improved by a median 8.2 km compared to 41% deteriorated by a median 6.2 km.
- 46% of the events are improved by 20% or more compared to 30% deteriorated by 20% or more.
- 90% error ellipse coverage reduced from 98% to 97%.

Again, the mislocation distribution for this test set of 318 events limited to the IMS and surrogate stations is not distinguishable from the larger set of 571 events using all stations. The mislocation 80th percentile is reduced 30% (from 53 to 37 km), 18% (19 vs. 16) more events are located within GTX accuracy, and 15% (187 vs. 162) more events are located within 18 km. These test results demonstrate SSSCs should improve locations based on regional data for a fuller IMS network Calibration does more improvement and "does no harm".

Mixing calibrated Pn and Sn phases with uncalibrated Pn and Sn phases

We next evaluate the effect of mixing calibrated and uncalibrated Pn and Sn phases in event location. The GT0-GT10 events are located using all stations in the study region, with and without SSSCs for IMS stations and IMS+surrogate stations, and 246 and 340 events were relocated, respectively.

Using IMS stations only, 58% of events are improved by a median 4.9 km, compared to 42% events deteriorated by a median 3.8 km. 10 more (176 vs. 166) events are located within 18 km with SSSCs. At the 80th percentile there is a 23% reduction in mislocation (27 vs. 22 km). Using IMS + surrogate stations, 54% events are improved by a median 5.5 km, compared to 46% events deteriorated by a median 5.1 km. 18 more (239 vs. 221) events are located within 18 km with SSSCs. At the 80th percentile there is a 21% reduction in mislocation (29 vs. 24 km)

Compared to large set of 571 events using only calibrated data, the location improvement and, in particular, deterioration are smaller. Large improvement occurred in the mislocation tails. In both cases the median error ellipse areas are reduced with similar 90% coverage (97% without SSSCs vs. 96% with SSSCs). The test results indicate that mixing calibrated and uncalibrated regional phases in the IMS network does no harm, and the improvements in calibration are only generally diluted.

Mixing calibrated Pn and Sn phases with uncalibrated Pg, Lg, and teleseismic phases

This test combines calibrated Pn and Sn phases with uncalibrated Pg, Lg, and teleseismic phases. This simulates the IMS network with the CUB 1.0 Model Pn and Sn SSSCs installed. Using only IMS stations and IMS+surrogate stations, 245 and 328 events, respectively, were located with and without SSSCs.

Using IMS stations only, 60% of the events are improved by a median 5.5 km and 40% of the events deteriorated by a median 8.6 km. The improvement is slightly smaller and deterioration is slightly larger than the case without uncalibrated data. Using IMS + surrogate stations, the location improvement is also ambiguous due to the large number of teleseismic phases dominating event location. In both cases the median 90% ellipse coverages are 96%. There are more events

located within the GT accuracy with SSSCs than without (19 vs. 8 for the former, and 16 vs. 12 for the latter). The test results show that the Pn and Sn SSSCs at least “do no harm” in event location when mixed with uncalibrated regional and teleseismic phase in a simulated IMS network. However, the test results from the full set of events did not simulate the situation of a small event detected by the IMS network.

We wished to use a subset of the data set to simulate the situation where teleseismic phases are not dominant. Small magnitude GT events would be ideal in such testing where there are a fair amount of both regional and teleseismic phases. Unfortunately magnitudes as well as signal-to-noise ratios are absent/inaccurate for most events in this data set. We used the teleseismic to regional phase ratio to select a subset of the events for further evaluation. The ratio histograms for the GT0-GT10 events are similar to those for the PIDC REB events during 1995/01/01-2000/12/31. In both cases a large number of events have a teleseismic to Pn and Sn ratios of about 3.

We selected events with regional-to-teleseismic ratios of 1 and 3 for testing. Using IMS stations and IMS + surrogate stations, 59 and 85 small events, respectively, were relocated with and without SSSCs. Using IMS stations only, over 80% of the events are improved by a median of 6 km, while less than 20% of the events are deteriorated by a median of 10 km. Using IMS + surrogate stations, 71% of the events are improved by a median 5-7 km, compared to 29% events are deteriorated by a median 1-3 km. This simulation argues that Pn and Sn SSSCs will improve locations for small events in the IMS network when mixed with uncalibrated regional and teleseismic phases.

Conclusions

In Phase 1 Pn and Sn SSSCs have been developed from ray tracing in the 3D CUB 1.0 Model for IMS stations in the Group-2 study region. The SSSCs and model errors have been successfully tested and validated both offline and online. The model based SSSCs and model errors were validated by comparing with JHD empirical path corrections. Improvement in event relocation and error ellipse (both size and coverage) were verified by relocating GT events in the Group-2 study region, with and without SSSCs. A large number of events were relocated and evaluated, including 425 Fennoscandian GT events, 571 high-quality Group-2 GT0-GT10 events, 15 MORT GT10 events, and 5955 GSETT-3 REB events.

Comparisons of JHD corrections with the CUB1.0 model show high correlation, particularly for paths longer than 5°. There is a 44% variance reduction of travel times, where the bulk standard error of JHD-CUB1.0 was 1.15 sec, compared to 1.53 sec from JHD-IASPEI91. Generally the JHD corrections seem to span a larger range than the CUB1.0 corrections, which is likely an effect of smoothing of the CUB1.0 model (2 by 2^o model). The estimated uncertainties of the CUB1.0 model also appear large compared with the scatter in the JHD corrections.

Relocation results using SSSCs show overall improvement in event locations and error ellipses. When 571 high-quality GT0-GT10 events are relocated using Pn and Sn SSSCs, 60% of the events are improved with a median improvement in mislocation of 8 km (10 km for GT0-GT2 events). All events have reduced error ellipses without losing 90% coverage. The median reduction in ellipse area is 2360 km^2 (from 4600 km^2 without SSSCs to 2240 km^2 with SSSCs). The CUB 1.0 Model based SSSCs and model errors performed well w.r.t. IASPEI91, and the results

met or exceeded the evaluation criteria of the 1999 Oslo recommendations. Relocation tests using the MORT GT10 events show that, with SSSCs, event locations are better correlated with the ridges and transforms. Testing of the real world situation using over 6000 GSETT-3 REB events reveals tighter event clusters with SSSCs and more events were located with SSSCs. A benchmark test relocating GT events previously used to test Fennoscandian 1D SSSCs indicates that the existing 1D SSSCs can be replaced with 3D SSSCs in Fennoscandia.

To simulate regional location with calibrated Pn and Sn phases in an IMS network GT0-GT10 events were relocated using only IMS stations and IMS + surrogate stations. Relocation results show that calibration does more improvement than no-calibration. Mixing calibrated and uncalibrated regional phases in the IMS network “does no harm” to event location. Testing using regional data combined with uncalibrated teleseismic phases indicates that event location is improved by Pn and Sn SSSCs for simulated IMS networks even with large numbers of teleseismic phases. The simulation results argue that Pn and Sn SSSCs will improve locations for small events in the IMS network when mixed with uncalibrated regional and teleseismic phases.

Compared to previous studies, the relocation results here are similar to those for Fennoscandia (Yang et al., 1999), and both are better than those for North America (Yang et al., 2000). Note that the current results have only used Pn and Sn SSSCs while Pg and Lg were also used for the two other regions. Further, our statistics are based on all events, including events along the border of the study region which may be poorly located since we only use regional phases from stations within the study region.

Improvements on SSSCs and validation testing data are expected in Phase 2. More details may be found at http://g2calibration.cmr.gov. Since the median improvement in event location is less than 10 km using SSSCs, it is important to assess the GT events as accurately as possible. In Phase 2 more data will be collected for better coverage and quality. In general, it is undesirable to mix calibrated and uncalibrated data in event location. Developing Pg and Lg SSSCs has a high priority. Teleseismic phase calibration may be explored since teleseismics play a dominant role in IMS event location. In Phase 2 we will also experiment with depth-dependent SSSCs and more realistic modeling errors that are closely coupled with the 3-D model and stations.

Table A-1. Event clusters (Appendix 7)

Cluster	Epicenter	GT	No Event	No Sta	Geometry	
			Fix/Free[1]	<20/>20[2]	Diam[3]	Var[4]
Algeria, Blida	36.68 2.49	25	1/24	117/19	37.5/25.1	5.5/4.5
Algeria, Bordjbou	36.60 4.77	25	1/4	93/16	22.1/17.2	6.9/4.6
Algeria, El Asnam	36.31 1.59	25	1/59	172/185	55.7/53.0	5.9/5.6
Algeria, Mascara	35.52 -0.11	25	1/19	51/0	32.1/36.7	5.6/5.0
Aqaba Central	28.81 34.65	5	1/11	89/132	34.1/31.7	4.8/4.9
Aqaba North	29.19 34.72	5	1/13	56/102	42.5/53.2	7.6/7.2
Aqaba South	28.48 34.76	5	1/15	91/142	37.3/41.9	5.5/7.4
France, Annecy	45.94 6.09	1	1/11	73/0	11.2/11.3	2.1/2.6
Georgia, Racha East	42.44 44.00	5	1/23	48/177	48.3/38.4	7.9/5.3
Georgia, Racha West	42.58 43.25	5	1/24	28/123	38.1/38.7	5.4/5.2
Greece, Alani	40.11 21.62	25	1/19	42/2	35.7/33.4	6.8/5.6
Greece, Amfissa	38.30 22.45	25	1/20	140/48	30.1/32.6	5.5/5.3
Greece, Crete	34.95 23.05	25	1/28	139/108	34.6/31.4	5.2/4.9
Greece, Ionian Sea	35.97 21.95	25	1/17	139/69	45.9/43.2	8.5/8.5
Greece, Kanallkion	39.27 20.55	25	1/17	196/52	21.2/32.3	3.7/5.7
Greece, Kefallnia	38.07 20.25	25	1/12	163/104	19.2/38.5	2.9/6.9
Greece, Pagasae	39.26 22.81	25	1/19	148/101	17.6/18.8	3.3/3.0
Greece, Thermum	38.52 21.67	25	1/25	147/56	25.0/26.1	4.0/4.1
Greece, Thivai	38.24 23.26	25	1/30	139/89	19.7/21.4	2.8/3.0
Greece, Zakynthos	37.63 20.94	25	1/18	147/37	24.2/22.1	3.5/3.3
Greece, Zmfissa	42.04 19.05	25	1/22	127/7	22.2/26.0	3.6/4.5
Italy, Abruzzo	41.76 13.90	25	1/25	140/36	21.5/18.4	3.3/2.6
Italy, Forli	44.24 12.21	25	1/22	94/0	26.8/29.3	5.6/5.3
Italy, Gemona	46.30 13.19	25	1/32	178/104	25.4/29.1	3.5/4.5
Italy, Reggio	44.69 10.32	25	1/18	102/1	39.4/38.1	5.6/5.8
Italy, Rionero Central	40.78 15.77	25	1/25	157/32	36.8/36.2	5.9/5.6
Italy, Rionero North	40.91 15.37	25	1/29	102/40	26.3/29.3	3.9/3.8
Italy, Rionero South	40.65 15.40	25	1/18	93/31	26.2/26.3	4.1/3.7
Italy, Taormina	37.90 15.06	25	1/18	79/7	35.1/29.4	5.7/4.9
Italy, Umbria-Marche	43.01 12.80	5	1/64	259/102	47.7/45.2	6.4/6.2
Italy, Ustica	38.44 12.78	25	1/19	114/32	32.0/26.2	5.5/4.1
Montenegro, Kotai	42.29 18.68	25	1/14	84/5	20.8/24.2	3.4/3.0
Morocco, Alhoceima	35.19 -4.04	5	2/21	61/8	23.4/18.0	3.5/2.1
Morocco, Melilla	35.35 -2.50	25	1/21	56/5	29.2/25.5	4.4/3.5
Poland, Lubin	51.49 16.09	1	9/37	131/9	36.6/15.6	4.7/2.1
Poland, Silesia	50.35 18.82	1	1/23	31/1	24.4/23.9	3.8/3.8
Russia, Astrakhan	46.76 48.27	1	1/14	27/156	24.1/20.9	5.0/4.9
Russia, Azgir	47.90 48.16	1	1/7	27/167	14.3/8.7	4.0/1.6
Russia, Orenburg	51.36 53.31	1	1/5	12/135	5.5/4.2	1.1/0.7
Slovenia, Krm Mountains	46.31 13.63	1	1/21	115/0	18.2/12.1	2.5/2.3
Spain, Jayena	36.96 -3.78	5	1/18	59/2	43.1/37.6	4.7/4.1
Spain, Loja	37.21 -4.20	5	4/29	38/0	31.3/15.7	4.4/1.9
Spain, Murcia	38.12 -1.48	5	2/7	23/0	29.8/34.2	6.2/5.0
Switzerland, Engelberg	46.72 8.42	0	1/7	13/0	31.0/31.2	8.4/6.6
Turkey, Adana	36.88 35.50	5	2/14	65/42	41.3/52.7	6.8/9.6
Yemen, Gulf of Aden North	14.05 51.65	25	1/11	2/92	49.6/44.9	8.3/7.9
Yemen, Gulf of Aden South	13.09 50.96	25	1/14	2/102	56.5/46.2	10.0/6.2

1. Number of events held fixed and free in the JHD
2. Number of stations with distances less than and greater than 20 degrees
3. Diameter of event cluster before and after JHD (km)
4. Scatter in event locations from center of cluster before and after JHD (km)

Tabele A-2. Summary of statistics for JHD and CUB1.0 corrections (Appendix 7)

Cluster	No. Obs	Median path (°)	Range		Standard Deviation (sec)			Correlation	
			JHD	CUB1.0	JHD	CUB1.0	JHD-CUB1.0	coeff.	p-value[1]
Algeria, Blida	108	8.4	-2.19 2.75	-2.00 1.16	0.86	0.44	0.92	0.32	0.00
Algeria, Bordjbou	78	8.1	-2.95 2.54	-1.38 1.03	1.11	0.50	1.05	-0.10	0.82
Algeria, El Asnam	136	10.9	-2.70 7.53	-2.51 1.01	1.34	0.87	1.45	0.09	0.16
Algeria, Mascara	51	6.4	-1.97 2.60	-2.11 1.82	0.89	0.79	0.80	0.57	0.00
Aqaba Central	57	4.3	-1.90 3.21	-1.88 1.22	0.61	0.44	0.70	0.07	0.30
Aqaba North	42	4.0	-2.09 7.52	-1.59 1.16	0.96	0.33	0.82	-0.06	0.68
Aqaba South	61	4.2	-1.39 5.37	-1.91 1.12	0.86	0.44	0.93	0.00	0.54
France, Annecy	58	2.2	-1.59 3.18	-0.97 0.90	1.45	0.53	1.05	0.67	0.00
Georgia, Racha East	38	13.5	-7.10 2.54	-4.74 2.09	1.92	1.04	1.35	0.75	0.00
Georgia, Racha West	20	13.2	-3.86 6.18	-4.06 2.76	1.96	1.87	1.16	0.37	0.05
Greece, Alani	40	3.0	-1.47 2.47	-0.76 0.46	1.23	0.24	1.19	0.11	0.25
Greece, Amfissa	105	4.9	-2.48 5.65	-1.56 0.85	1.38	0.44	1.48	-0.02	0.64
Greece, Crete	116	7.8	-3.61 7.38	-1.71 0.94	1.63	0.69	1.39	0.19	0.02
Greece, Ionian Sea	125	6.8	-3.13 8.27	-1.52 1.24	1.30	0.67	0.90	0.39	0.00
Greece, Kanallkion	183	6.5	-4.37 10.42	-2.81 1.27	1.38	0.50	1.19	0.44	0.00
Greece, Kefallnia	147	8.2	-5.08 5.50	-2.57 0.96	1.50	0.58	1.17	0.52	0.00
Greece, Pagasae	128	7.8	-3.11 4.88	-1.94 1.80	1.43	0.61	1.27	0.33	0.00
Greece, Thermum	130	5.4	-2.75 5.33	-2.19 1.20	1.86	0.56	1.62	0.03	0.38
Greece, Thivai	124	8.0	-3.21 7.36	-1.50 1.62	1.78	0.83	1.92	0.08	0.19
Greece, Zakynthos	135	5.1	-3.91 8.64	-1.91 1.42	1.23	0.44	0.99	0.07	0.22
Greece, Zmfissa	118	5.9	-3.19 4.10	-3.21 0.89	1.37	0.42	1.24	0.39	0.00
Italy, Abruzzo	129	7.1	-2.10 3.61	-0.87 0.69	1.08	0.46	0.98	0.22	0.01
Italy, Forli	89	3.5	-2.29 2.19	-0.80 0.61	0.70	0.30	0.80	0.03	0.38
Italy, Gemona	168	6.0	-2.35 6.42	-3.89 1.53	0.85	0.73	1.16	0.03	0.36
Italy, Reggio	102	2.9	-1.67 2.13	-1.05 1.06	0.83	0.65	0.80	0.35	0.00
Italy, Rionero Central	151	6.1	-4.36 7.38	-2.18 0.81	0.98	0.42	1.11	0.34	0.00
Italy, Rionero North	95	7.0	-3.31 3.15	-0.87 0.83	1.10	0.44	1.13	0.01	0.47
Italy, Rionero South	87	7.4	-3.58 14.48	-2.09 0.71	1.20	0.56	1.33	-0.01	0.59
Italy, Taormina	65	5.9	-3.01 3.52	-0.80 0.69	1.53	0.24	1.38	0.06	0.30
Italy, Umbria-Marche	248	6.1	-3.28 3.26	-1.85 1.01	1.18	0.39	1.18	0.23	0.00
Italy, Ustica	109	7.6	-2.88 6.09	-0.78 1.47	1.85	0.31	1.84	0.36	0.00
Montenegro, Kotai	77	6.7	-3.04 3.20	-3.06 1.09	0.99	0.64	1.11	0.38	0.00
Morroco, Alhoceima	55	3.4	-1.55 2.37	-1.53 0.69	0.85	0.43	0.89	0.11	0.21
Morroco, Melilla	53	5.5	-1.92 2.13	-1.35 1.32	0.95	0.37	0.89	0.25	0.04
Poland, Lubin	127	6.4	-5.80 3.67	-3.93 1.01	1.22	0.74	0.95	0.74	0.00
Poland, Silesia	27	5.0	-5.28 1.80	-4.34 1.31	1.35	1.32	0.64	0.80	0.00
Russia, Astrakhan	19	16.1	-2.37 7.36	-1.46 2.15	1.44	0.89	1.57	0.38	0.05
Russia, Azgir	16	15.0	-2.45 3.64	-2.52 2.37	2.44	1.60	0.87	0.87	0.00
Slovenia, Krm Mountains	104	4.2	-1.73 2.19	-0.94 0.89	0.99	0.56	1.10	-0.06	0.76
Spain, Jayena	58	5.9	-1.97 3.37	-1.34 1.58	0.70	0.87	0.86	0.32	0.01
Spain, Loja	31	3.4	-1.30 3.20	-0.89 1.56	0.82	0.67	0.83	0.32	0.04
Spain, Murcia	22	3.1	-1.52 2.08	-0.66 0.98	0.59	0.46	0.82	0.33	0.07
Turkey, Adana	54	5.2	-3.63 3.23	-0.85 1.02	1.05	0.24	0.58	0.19	0.08

1. p-value from Pearson's correlation test.

Table A-3. Major evaluation metrics of offline validation testing from relocation of all GT0-GT10 events using all regional Pn and Sn SSSCs (Appendix 8)

	All events (GT0-10)	**All events (GT0-5)**	**All events (GT0-2)**	**All events (GT5)**
GT category	GT0-GT10	GT0-GT5	GT0-GT2	GT5
time period	5/1/1962-9/5/ 2000	5/1/1962-7/18/ 2000	5/1/1962-7/18/ 2000	10/25/1964-4/21/ 2000
number of events	571	548	272	276
median mislocation (km)	from 16.5 to 14.1	from 15.9 to 13.6	from 17.7 to 15.0	from 13.2 to 11.9
20,40,60,80 percentiles without and with SSSCs (km)	8;13;22;43 6;11;17;29	8;12;20;41 6;10;16;28	8;14;25;49 6;11;17;30	8;12;17;27 6;10;15;26
% of events improved on GT distance median improvement	60% 7.9 km	60% 7.6 km	61% 9.6 km	59% 5.5 km
% of events deteriorated on GT distance median deterioration	40% 6.4 km	40% 6.3 km	39% 6.4 km	41% 5.9 km
% of events improved on GT distance by more than 20%	47%	47%	49%	46%
% of events deteriorated on GT distance by more than 20%	31%	31%	29%	33%
% of events improved on confidence ellipses median improvement (sqkm)	100% 2100 (from 4600 to 2240)	100% 2000 (from 4300 to 2150)	100% 2600 (from 5620 to 3000)	99% 1000 (from 1930 to 900)
90% ellipse coverage	from 97% to 91%	from 97% to 90%	from 97% to 93%	from 96% to 88%
20,50,80,90 percentiles without and with SSSCs	0.02;0.07;0.3;0.5 0.02;0.09;0.5;0.9	0.02;0.07;0.3;0.5 0.02;0.08;0.4;1.0	0.01;0.07;0.3;0.5 0.01;0.07;0.4;0.7	0.02;0.07;0.3;0.5 0.02;0.1;0.5;1.2
Trinity	from 11% to 24%	from 11% to 25%	from 2% to 9%	from 20% to 42%
% of events improved on origin time error median improvement	99% 0.9 s (from 3.1 to 2.2)	99% 0.7 (from 3.0 to 2.2)	99% 0.8 (from 3.8 to 2.7)	99% 0.6 s (from 1.8 to 1.2)
% of event improved on standard error of observations median improvement	61% 0.2 (overall from 1.2 to 1.1)	61% 0.2 (overall from 1.2 to 1.1)	-	71% 0.2 (overall from 1.3 to 1.1)

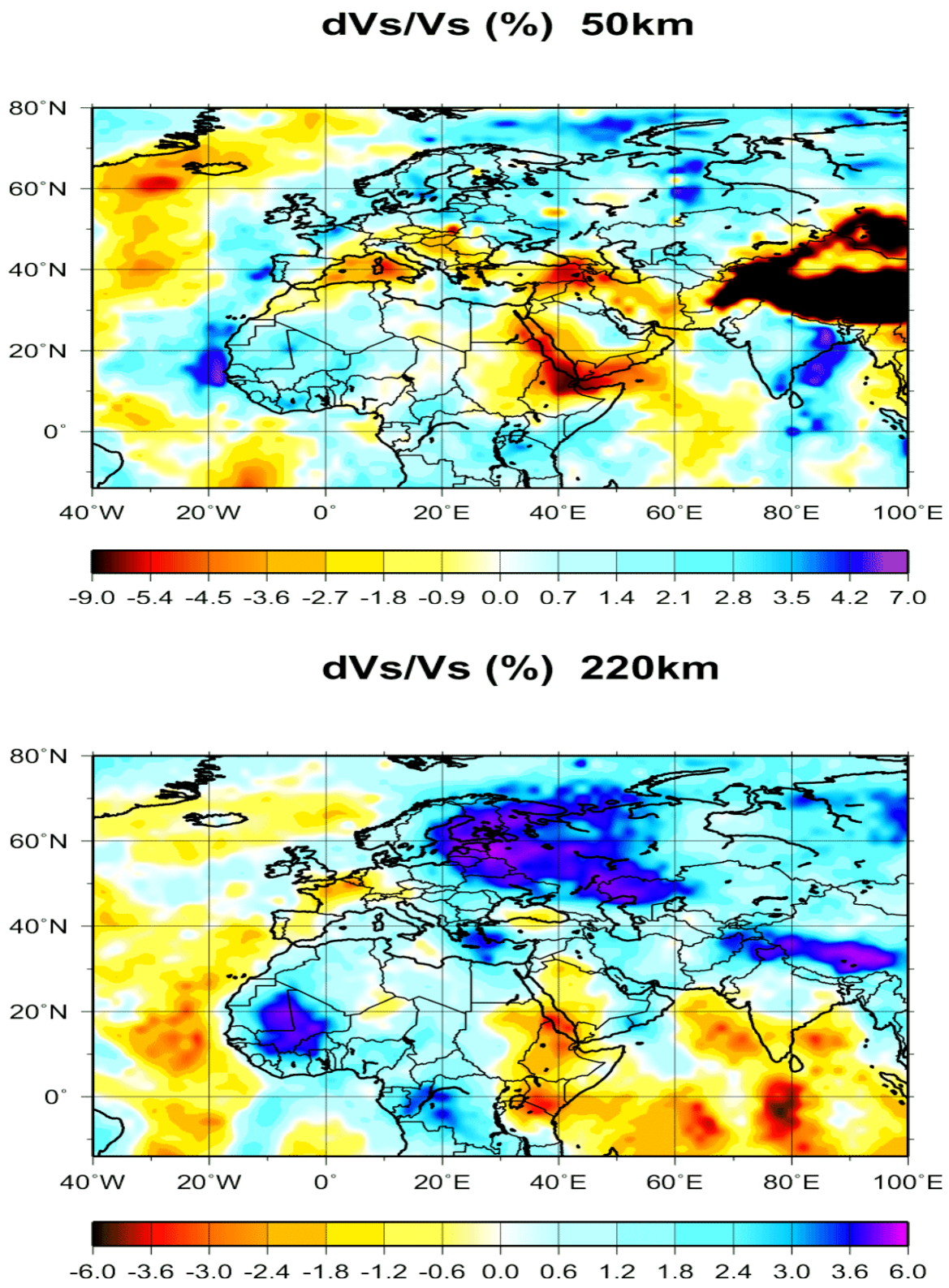

Figure A-1. CUB 1.0 Model. Shear wave velocity variations at depths of 50 km (upper) and 220 km (lower) (Appendix 2).

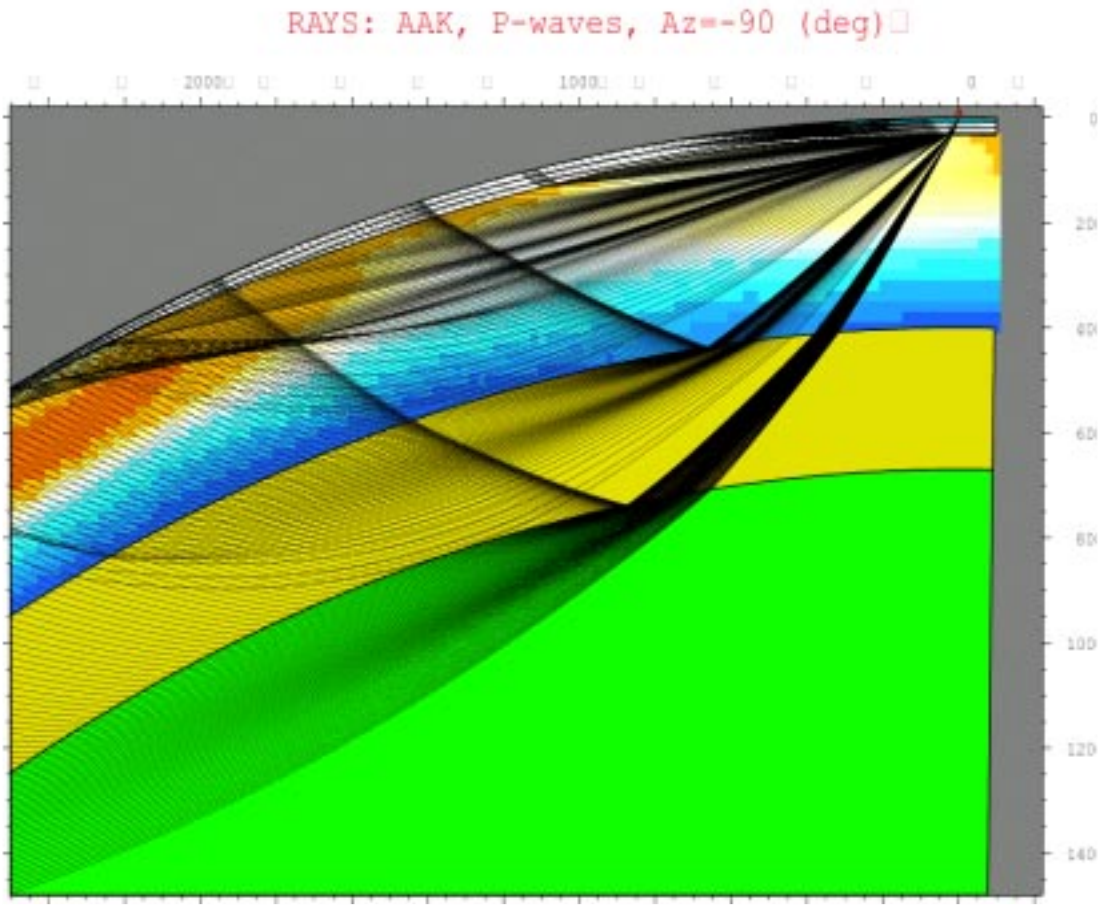

Figure A-2. Ray tracing of the CUB1.0 Model. P wave ray paths at azimuth of 270° are shown for station AAK (42.63° N, 74.48° E) (Appendix 3).

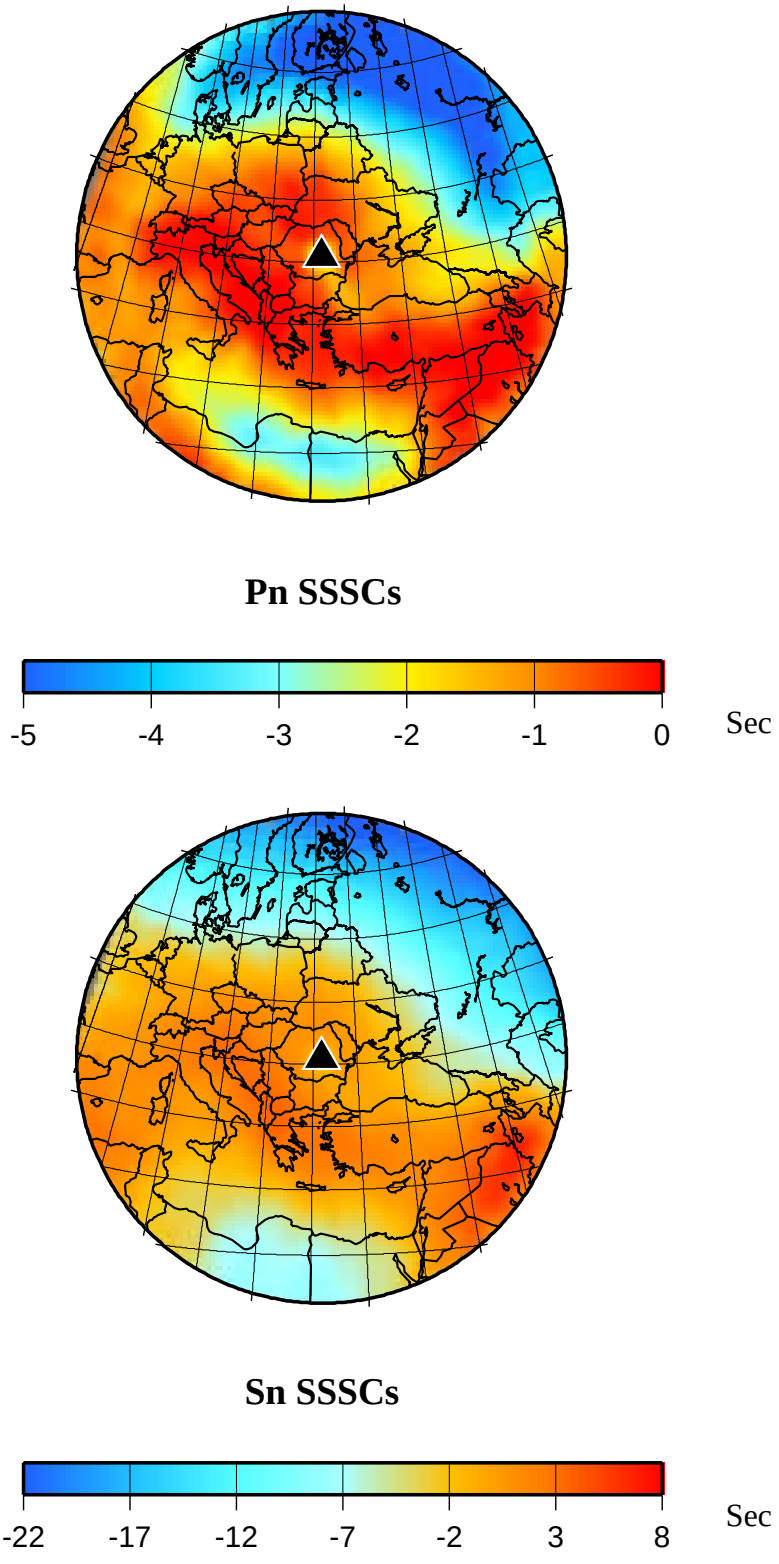

Figure A-3. An example of SSSCs calculated by ray tracing the 3D model (Appendix 3). Pn (top) and Sn (bottom) SSSCs for 3-component IMS station MLR (Muntele Rosu, Romania; triangle).

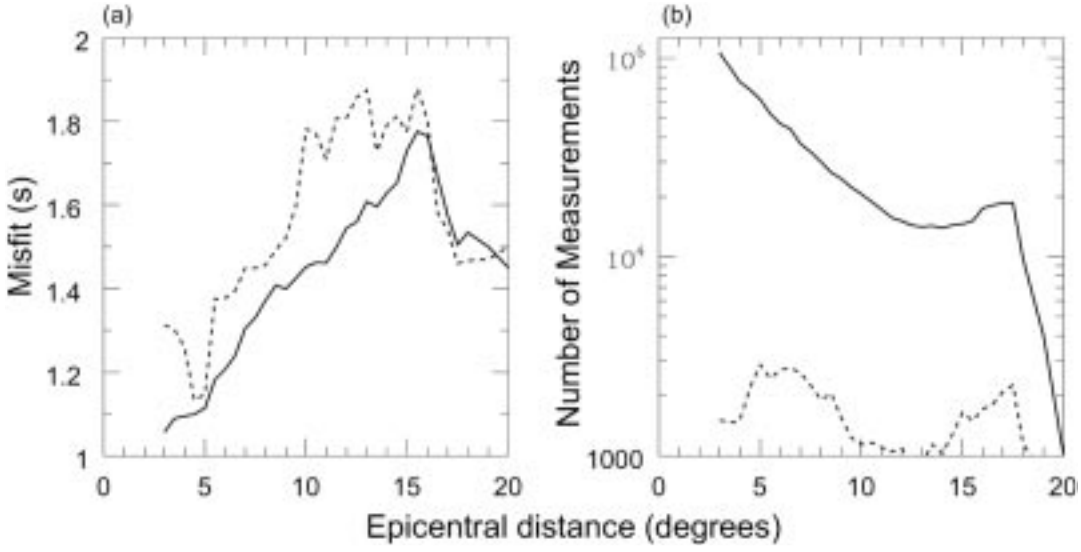

Figure A-4. (a) Average 1-norm misfit between Pn wave travel times and the predictions from the CUB1.0 Model as a function of epicentral distance. (b) Number of measurements in each 0.5º bin. The solid line shows data for all stations in the Group-2 region and the dashed line for only the IMS stations. Approximately 1,000,000 Pn arrivals were used. The Pn model error estimates are obtained from the variance of misfit for all stations (Appendix 4).

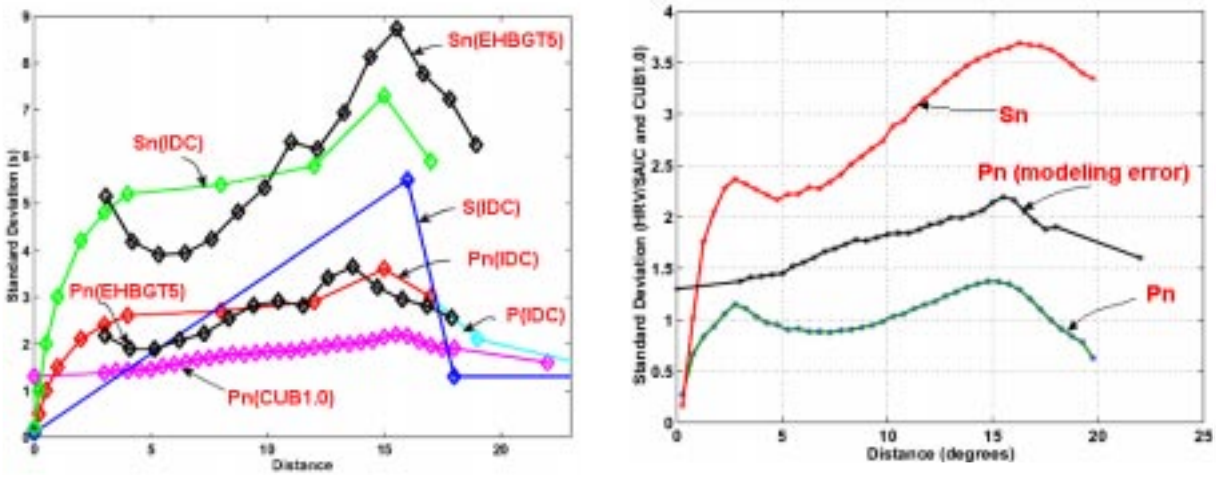

Figure A-5. Model errors used for Phase 1 and IDC, compared with those estimated from the EHB GT5 data set (left). They are the standard deviations of the residuals w.r.t. IASPEI91. (Appendix 4). (Right) Variation of travel time differences between the SAIC-HRV and CUB1.0 models, σ(SAIC/HRV - CUB1.0)/$\sqrt{2}$, compared to the Phase 1 Pn model errors. The error estimates obtained from the travel time differences are only part of the error budget.

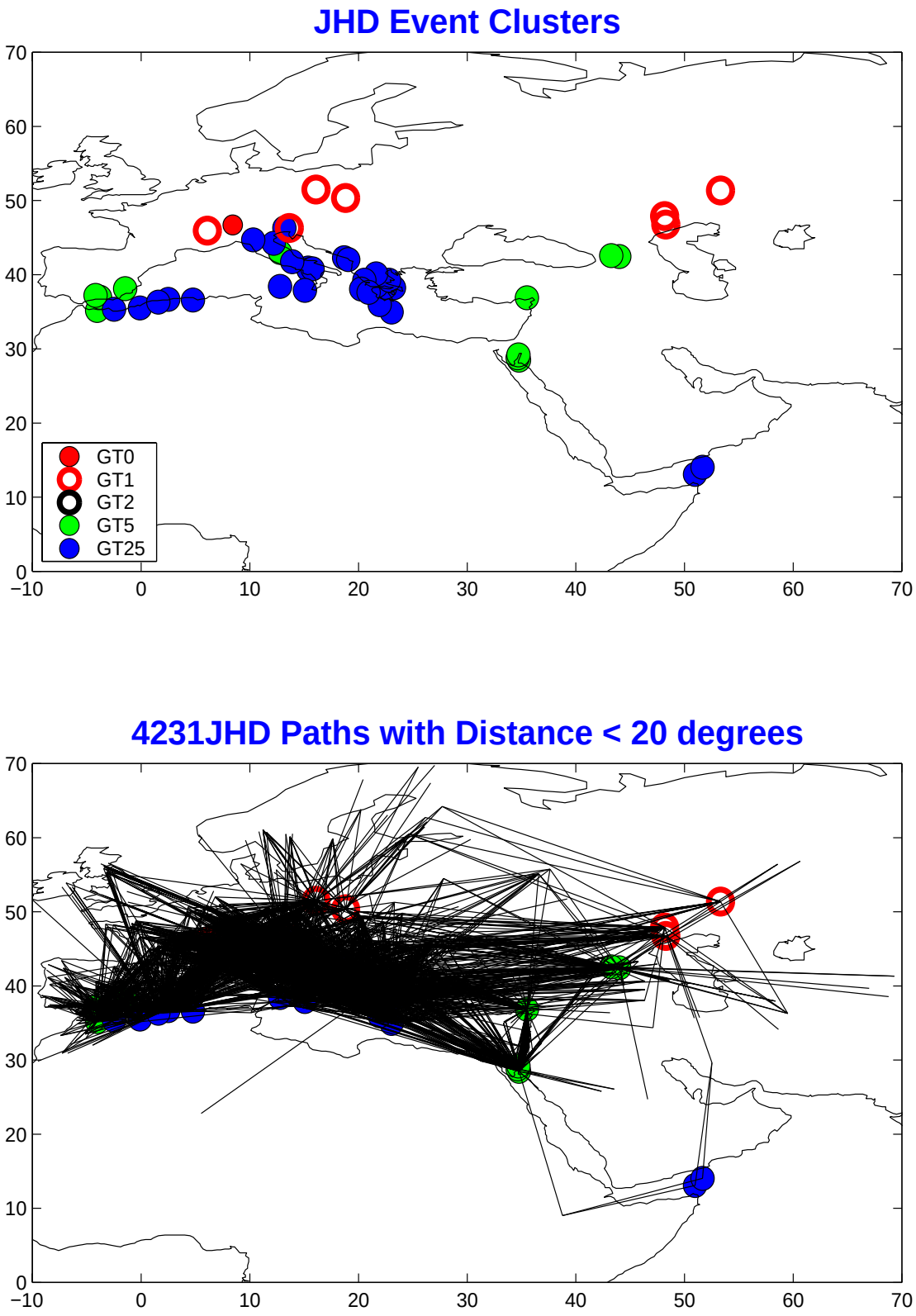

Figure A-6. Distribution of event clusters (top) and station path coverage (bottom) (Appendix 7).

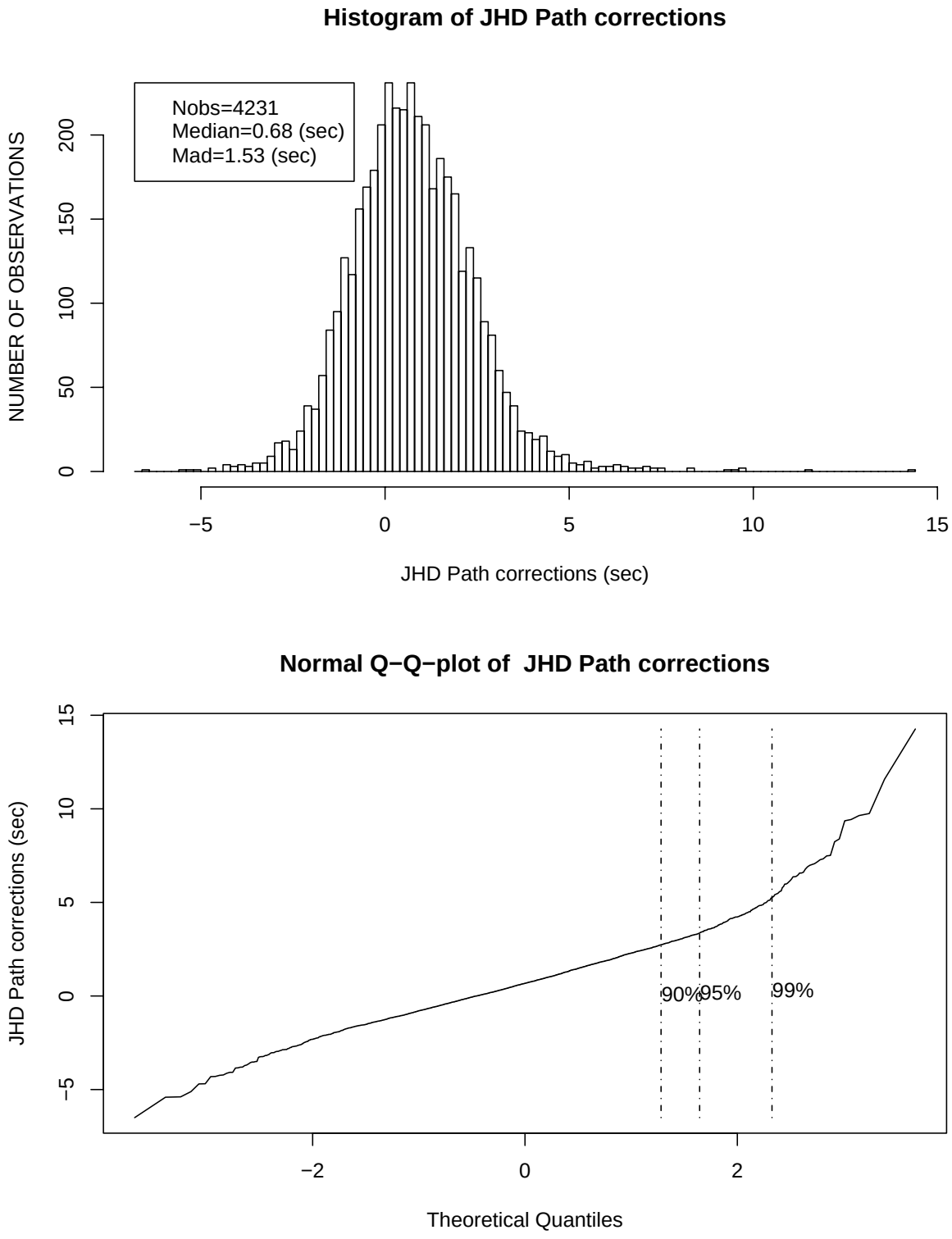

Figure A-7. Histogram of all JHD path correction illustrating their spread (top) and normal Q-Q plot (bottom) showing that the distribution is only approximately Gaussian. The scaling of the Q-Q plot makes empirical cumulative distributions for Gaussian follow a straight line (Appendix 7).

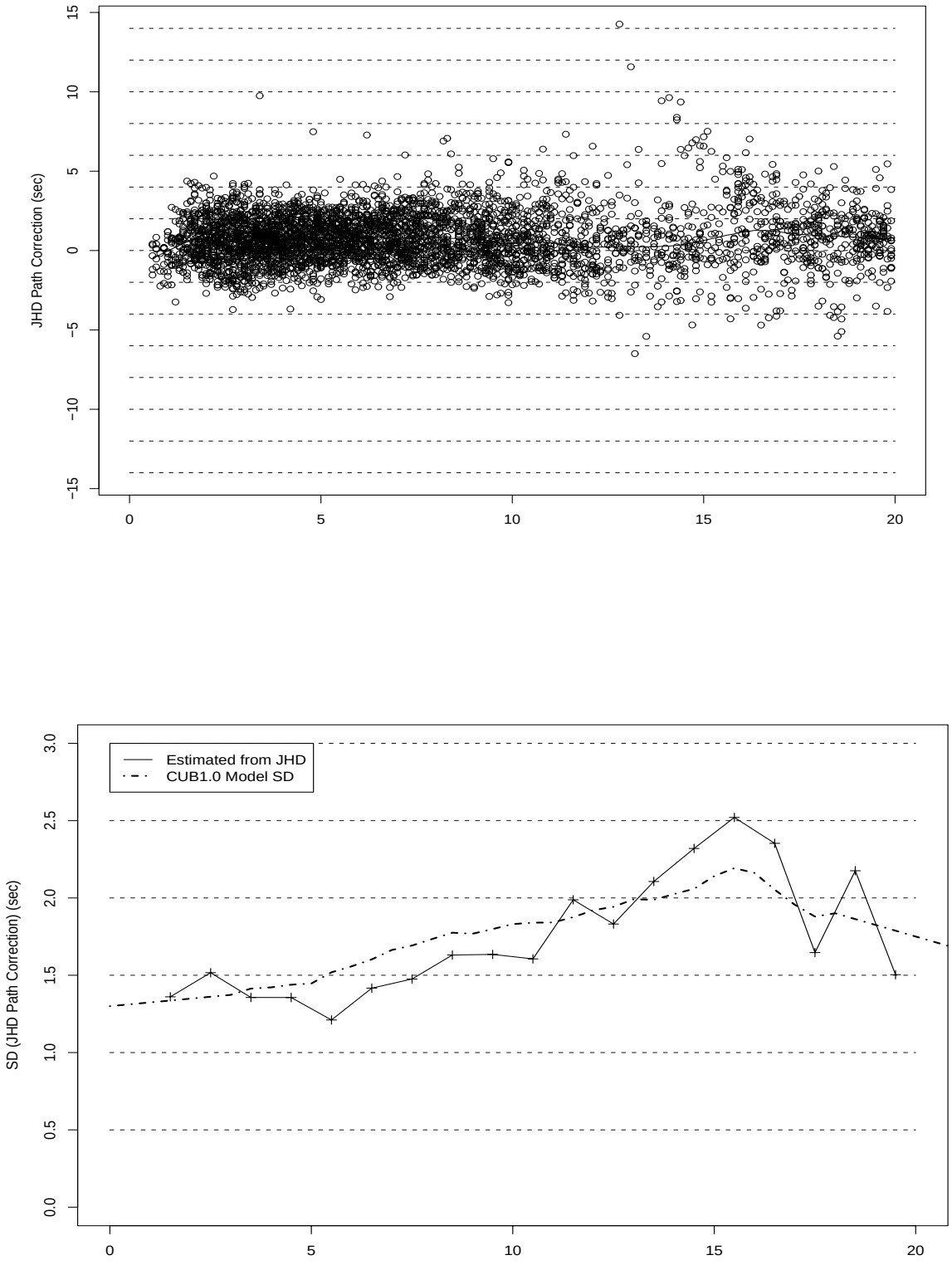

Figure A-8. JHD path corrections as a function of distance (upper) and the standard deviation of the path corrections as a function of distance (lower). The distance dependent a priori model error of the CUB1.0 model is drawn in the lower diagram for comparison (Appendix 7).

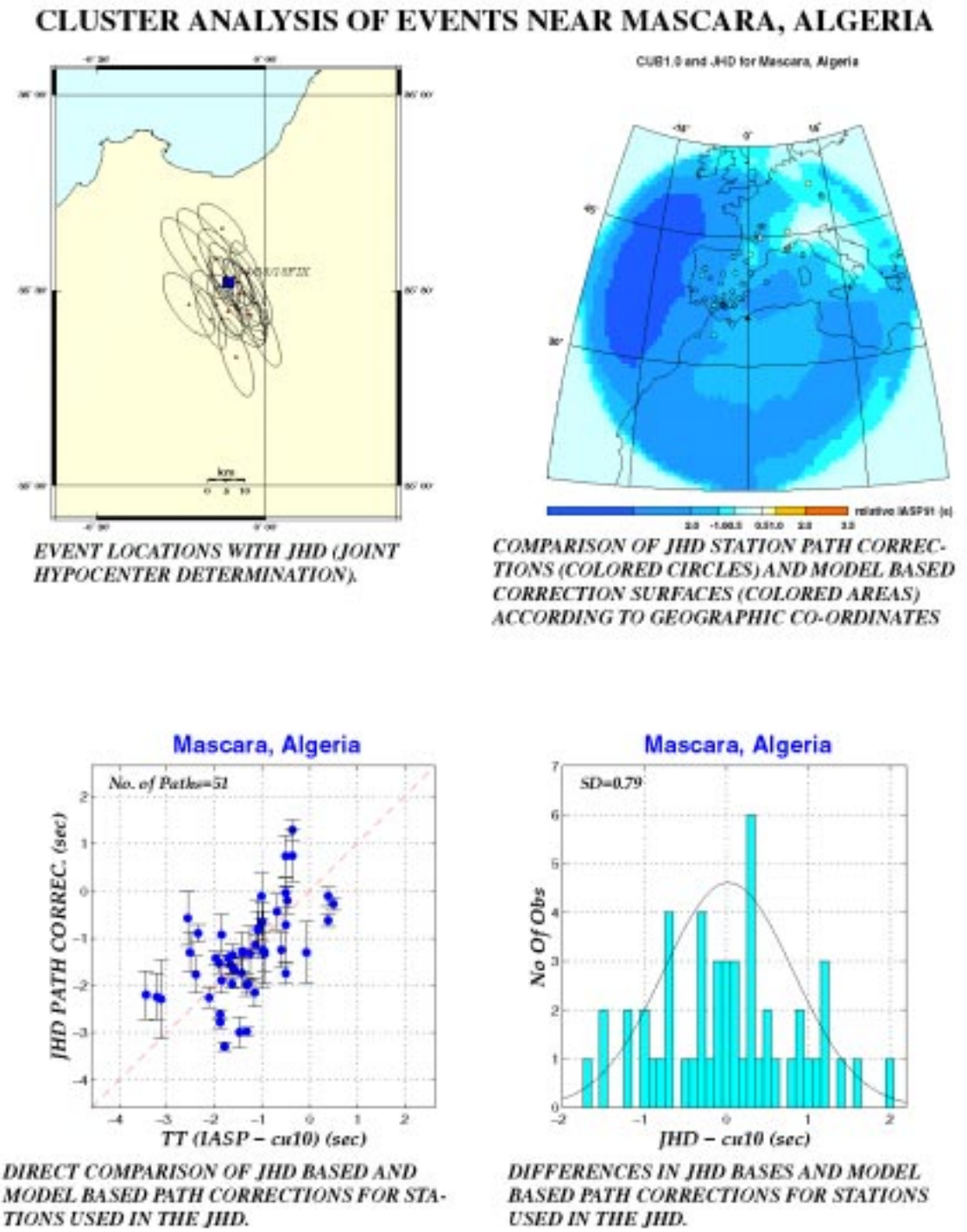

Figure A-9. Cluster analysis of events near Mascara, Algeria. Comparisons of JHD path corrections and SSSCs show good agreement. (Appendix 7).

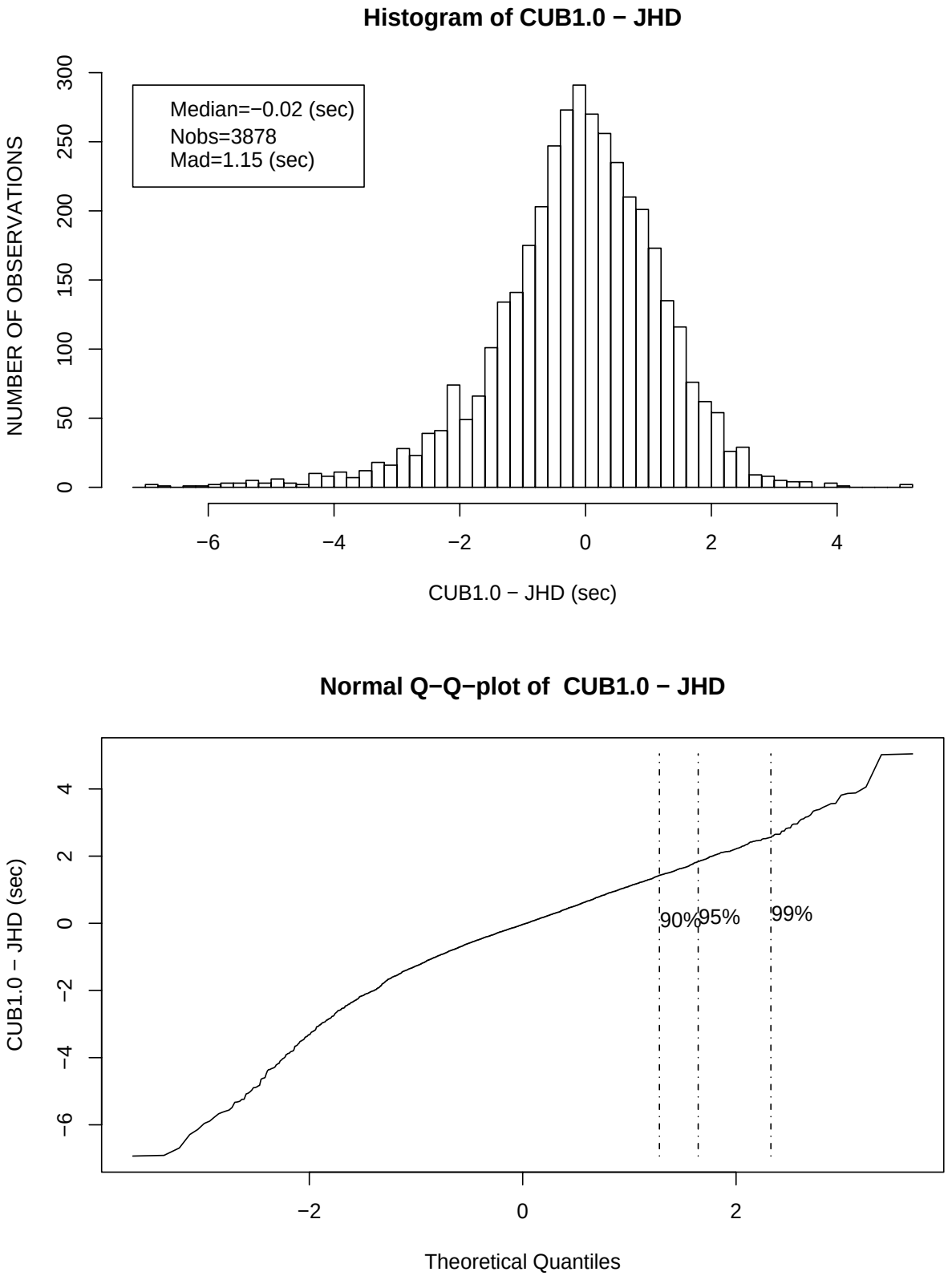

Figure A-10. The distribution of the CUB1.0-JHD differences (Appendix 7).

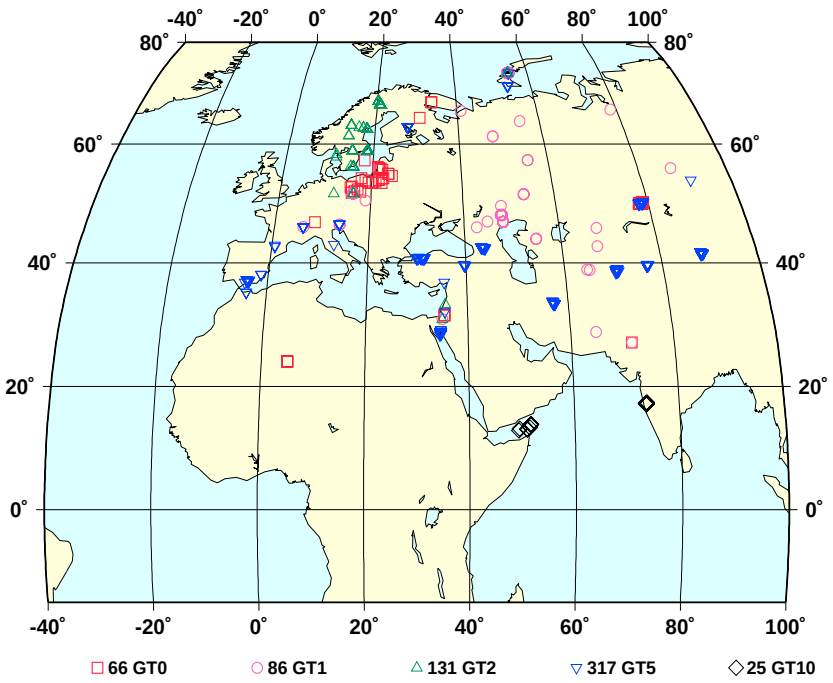

Figure A-11. Group-2 GT0-GT10 events used in offline validation testing (Appendices 8-9).

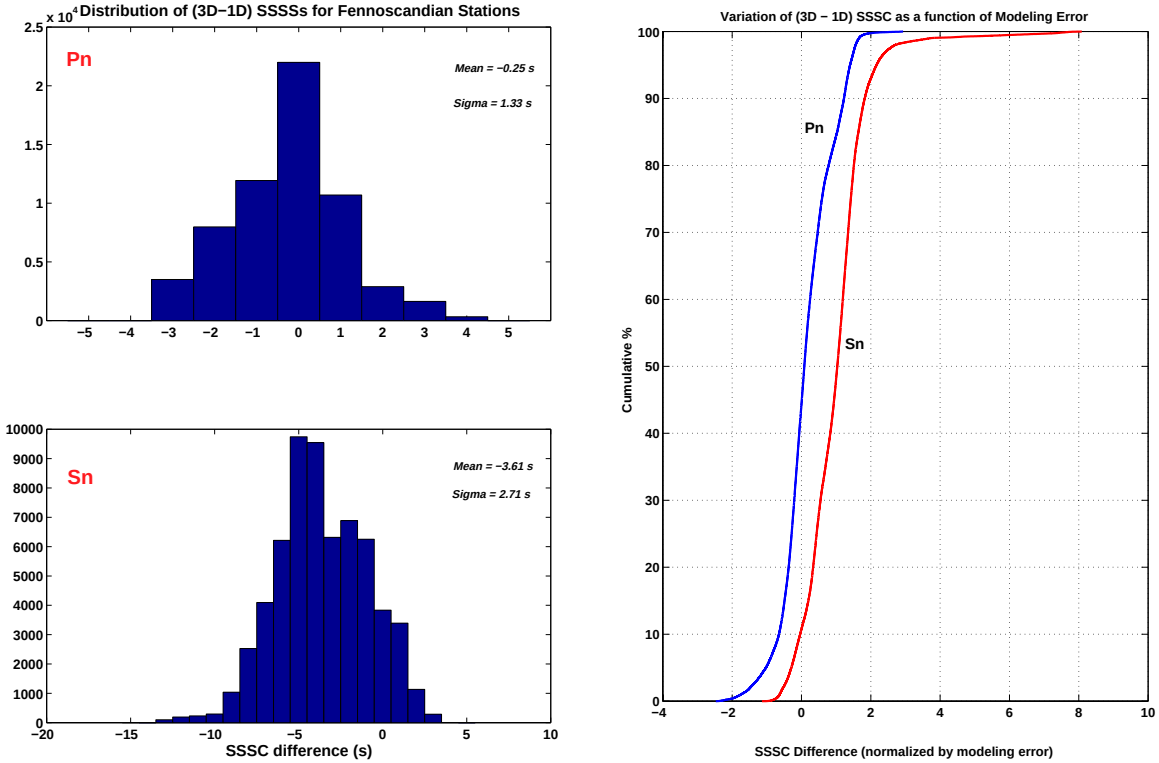

Figure A-12. Histograms and cumulative distributions of Pn and Sn differences between 1D and 3D SSSCs for Fennoscandia stations. The mean misfits are -0.25 sec for Pn and -3.61 sec for Sn, with the CUB model being faster. 84% and 100% of the Pn differences are below the 1σ and 2σ levels, respectively (48% and 93% for Sn), where σ is the corresponding CUB modeling error.

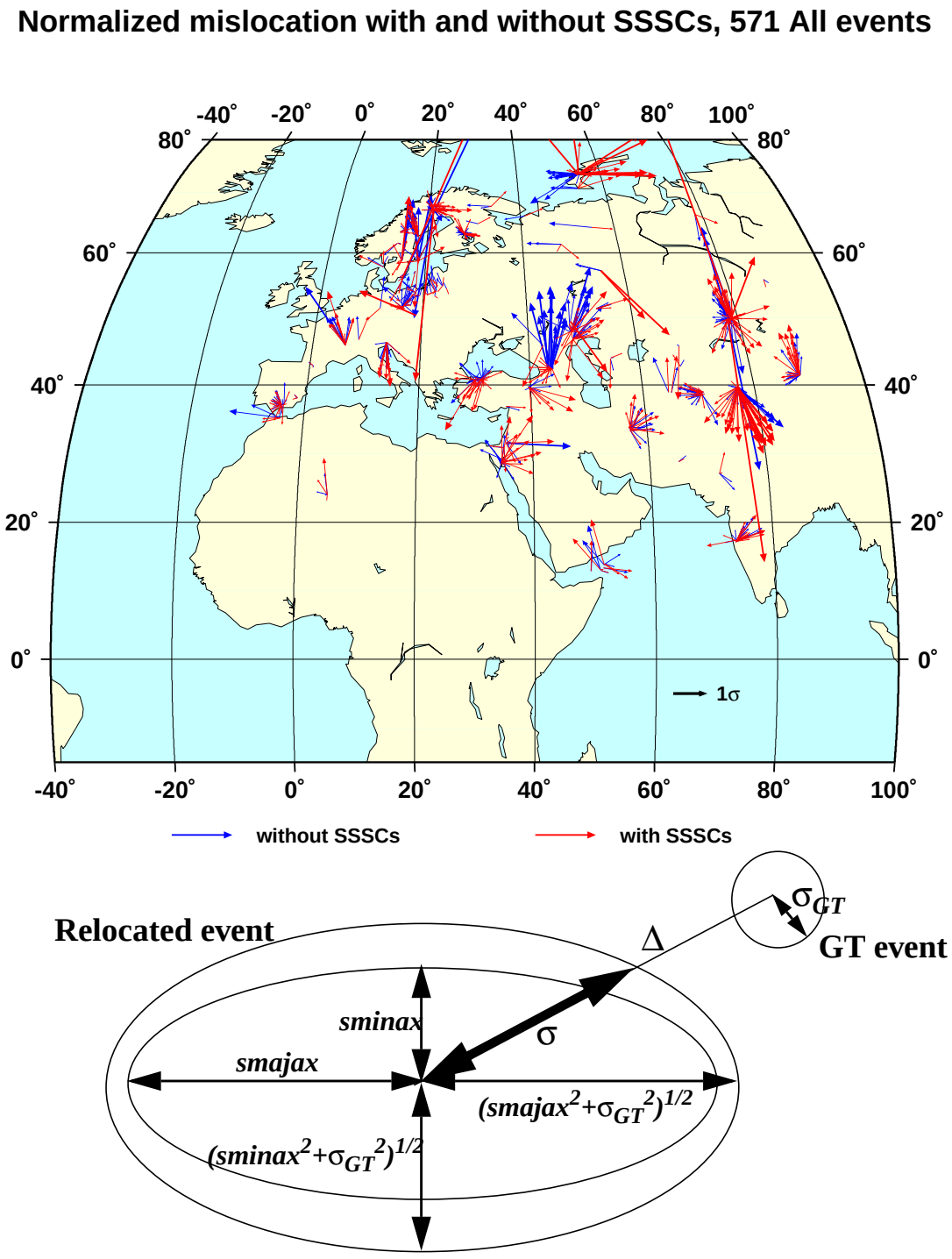

Figure A-13. Normalized mislocations (scaled by GT accuracy as shown in the bottom diagram) of GT0-GT10 events relocated with (red) and without (blue) SSSCs (Appendix 8). The direction of the arrows is from the GT to the relocation. The baseline scale is 1, i.e. the 90% coverage is met.

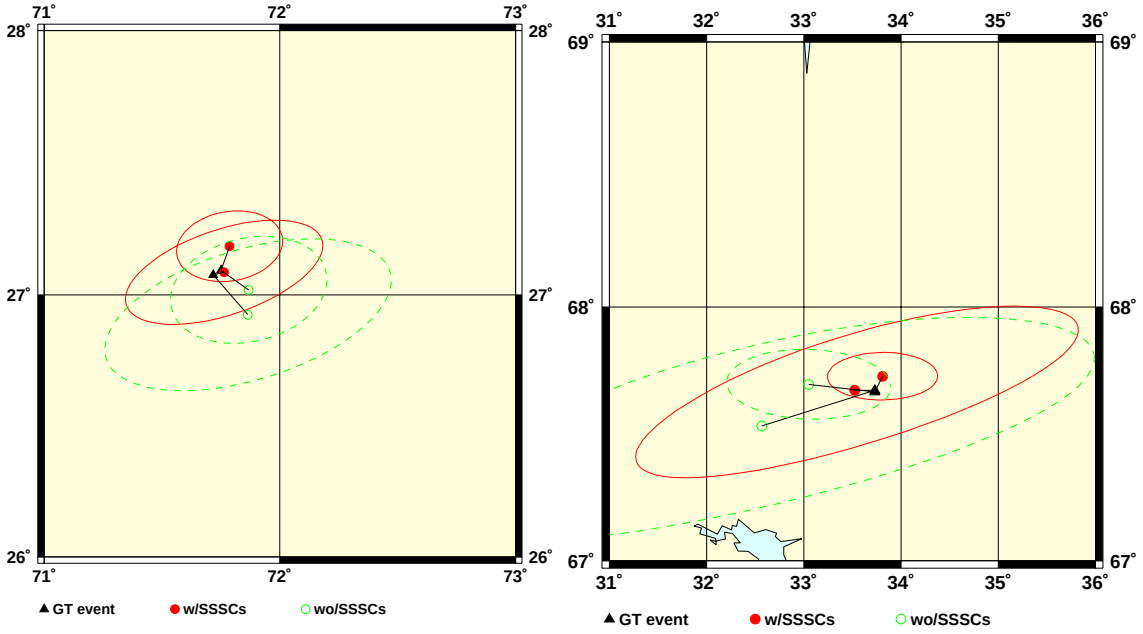

Figure A-14. Relocations of the 1974 and 1998 Indian nuclear explosions (Left; GT0) and the 1996/09/29&1997/10/12 Kola calibration shots (Right; GT0) using Pn and Sn phases, with (solid) and without (dashed) SSSCs (Appendix 8).Using SSSCs the mislocations are 4.6-10.8 km for the former, improved by 3.4 -17.6 km, and 7.6-8.8 km for the latter, improved by 21.4-43.2 km.

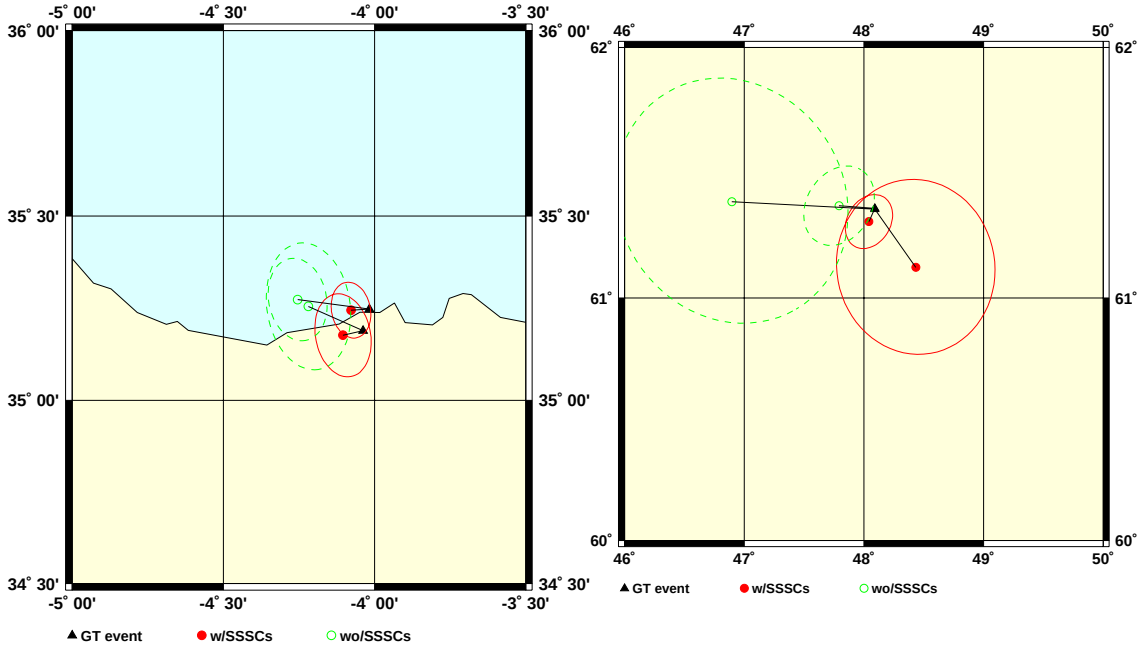

Figure A-15. Relocations of the Morocco cluster (Left; GT5) and two PNEs on 1971/10/04 and 1988/09/06 (Right; GT1) using Pn and Sn phases only, with (solid) and without (dashed) SSSCs (Appendix 8). Using SSSCs the mislocations are 5.6-6.2 km for the former, improved by 11.8-16.2 km, and 3.7-29.2 km for the latter, improved by 12.4-35 km.

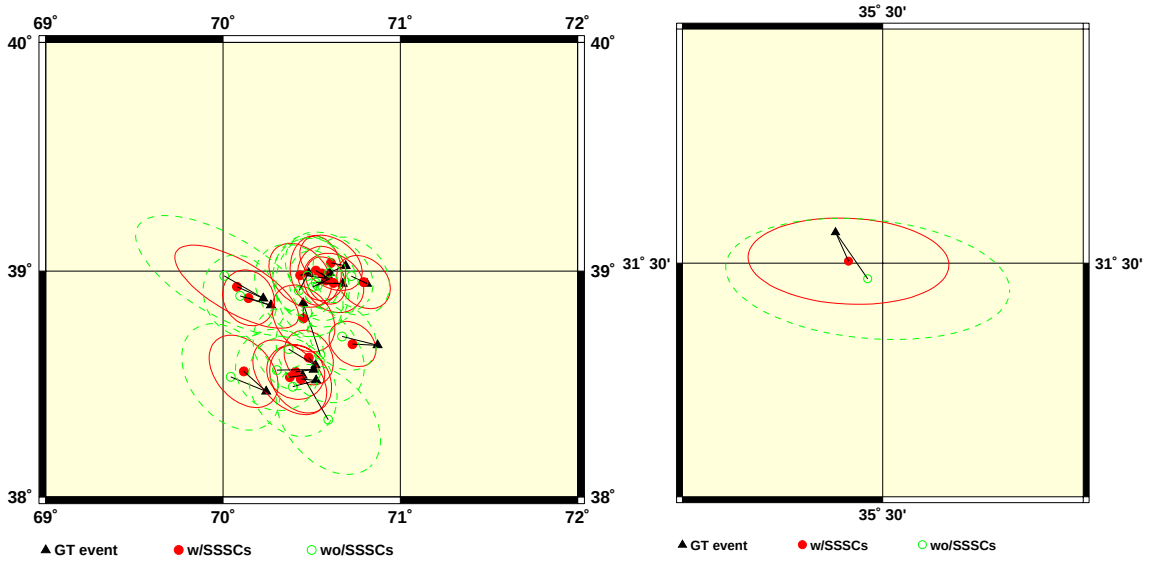

Figure A-16. Relocations of the Garm earthquake cluster (Left; GT5) and the 1999/11/11 Dead Sea shot (Right; GT0) using Pn and Sn phases only, with (solid) and without (dashed) SSSCs (Appendix 8). With SSSCs the median mislocation is 7.5 km for the former, improved by 5.9 km. The mislocations is 3.8 km for the latter, improved by 2.9 km.

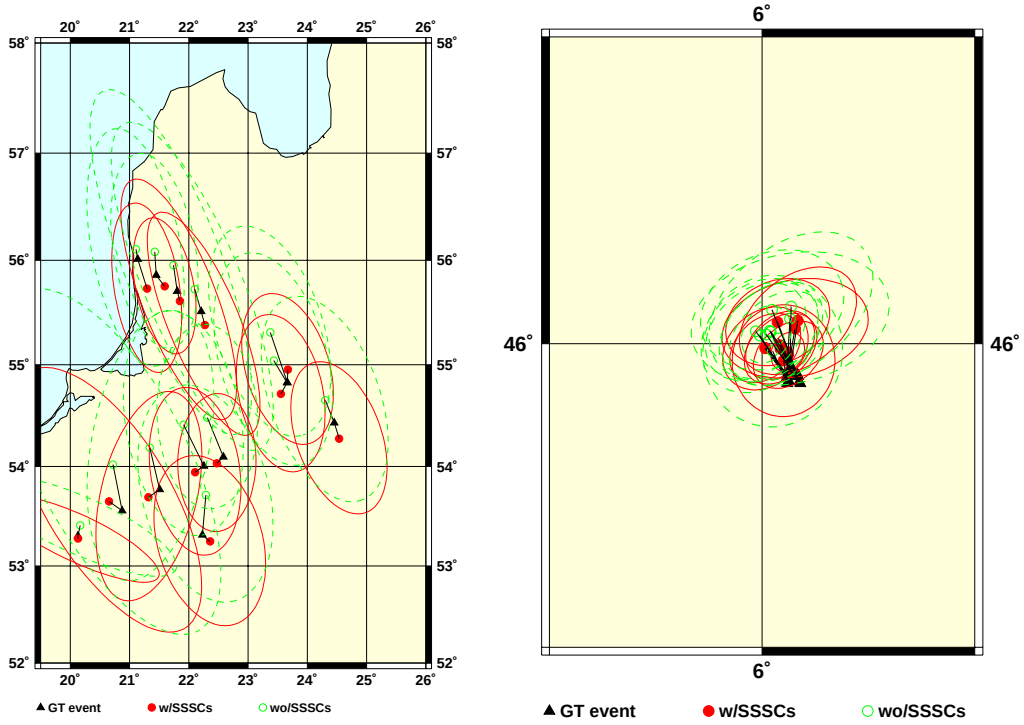

Figure A-17. Relocations of the 1997 Polonaise and 1995-1996 Eurobridge shots (Left; GT0) and the French earthquake cluster (Right; GT5) using Pn and Sn phases only, with (solid) and without (dashed) SSSCs (Appendix 8). Using SSSCs the median mislocation is 14 km for the former, improved by 24.7 km, and 6.2 km for the latter, improved by 2.8 km.

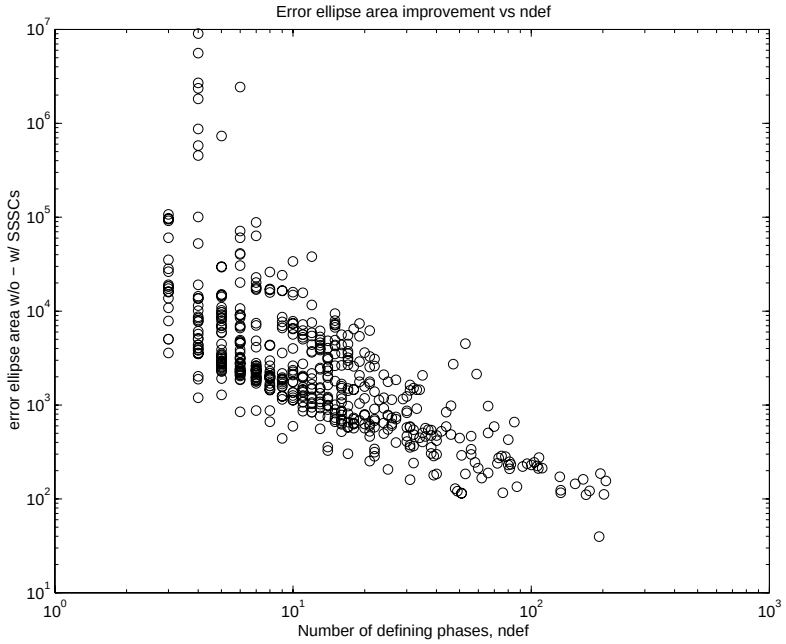

Figure A-18. Error ellipse area improvements for Group-2 GT0-GT10 events relocated with SSSCs, compared without SSSCs, as a function of number of defining phases (ndef). With SSSCs all events have reduced ellipse areas. There are large area reductions for low ndef events.

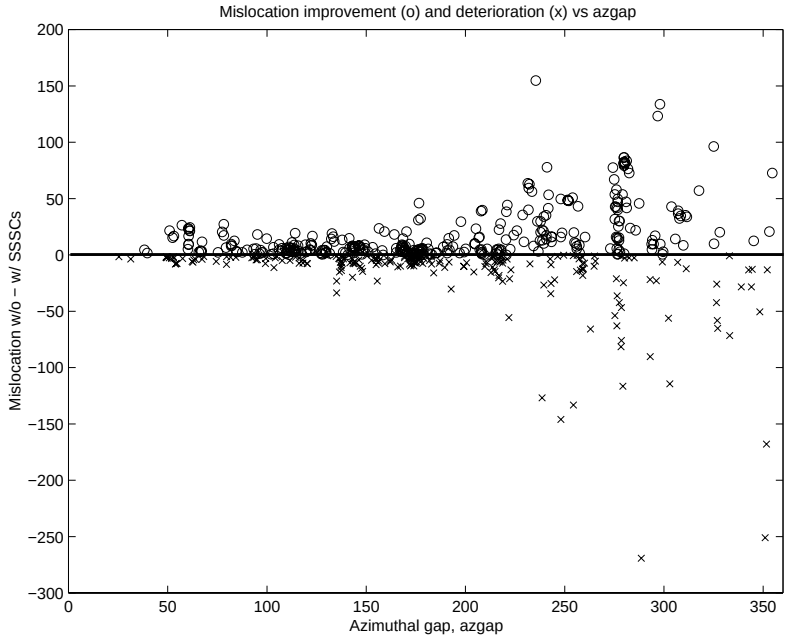

Figure A-19. Mislocation improvements (circle) and deteriorations (cross) for Group-2 GT0-GT10 events relocated with SSSCs, compared without SSSCs, as a function of azimuthal gap (azgap). Mislocations are similar with and without SSSCs up to azgap of 230 degrees.

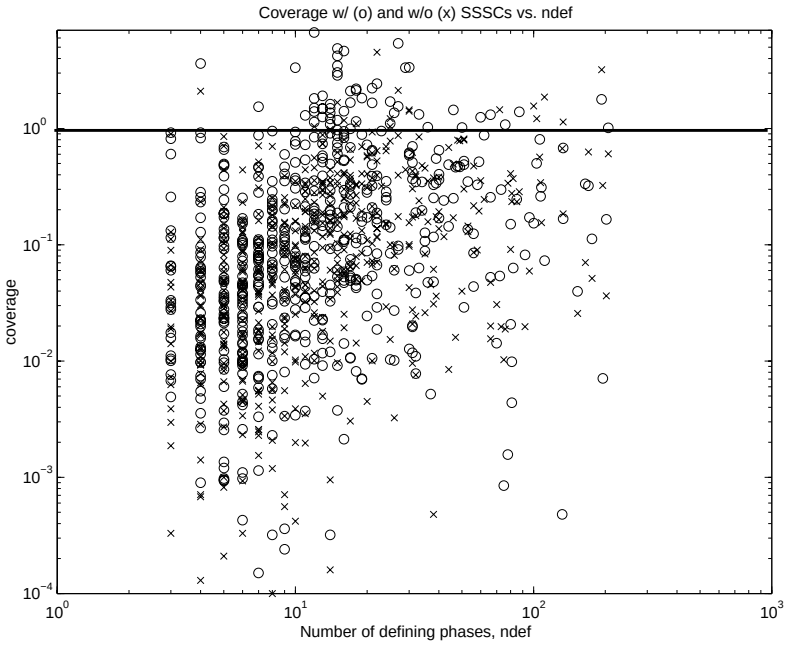

Figure A-20. 90% coverages for Group-2 GT0-GT10 events relocated with (circle) and without (cross) SSSCs. Symbols below the line (value of 1) indicates event relocations met 90% coverages. The 90% coverages are similar with and without SSSCs for low ndef events.

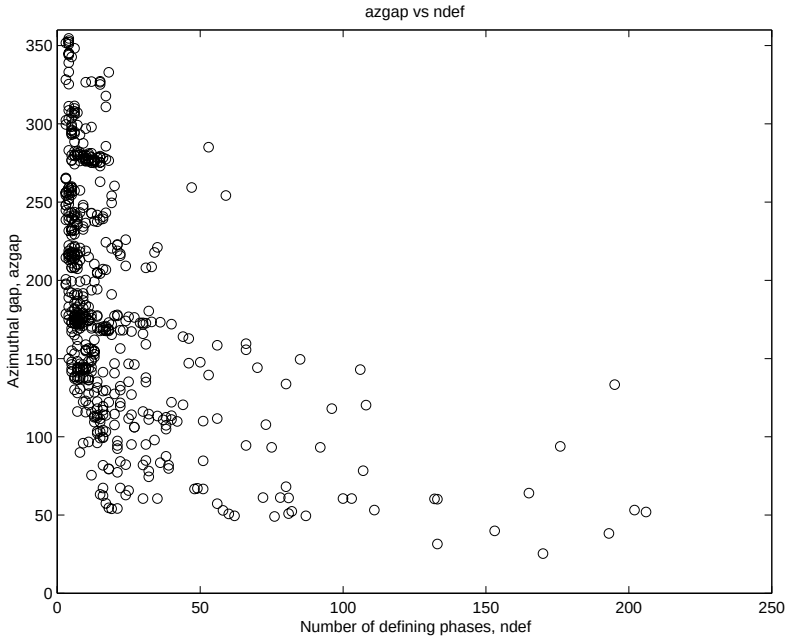

Figure A-21. Events with low ndef tend to have large azgap, but events with large ndef may also have large azgap.

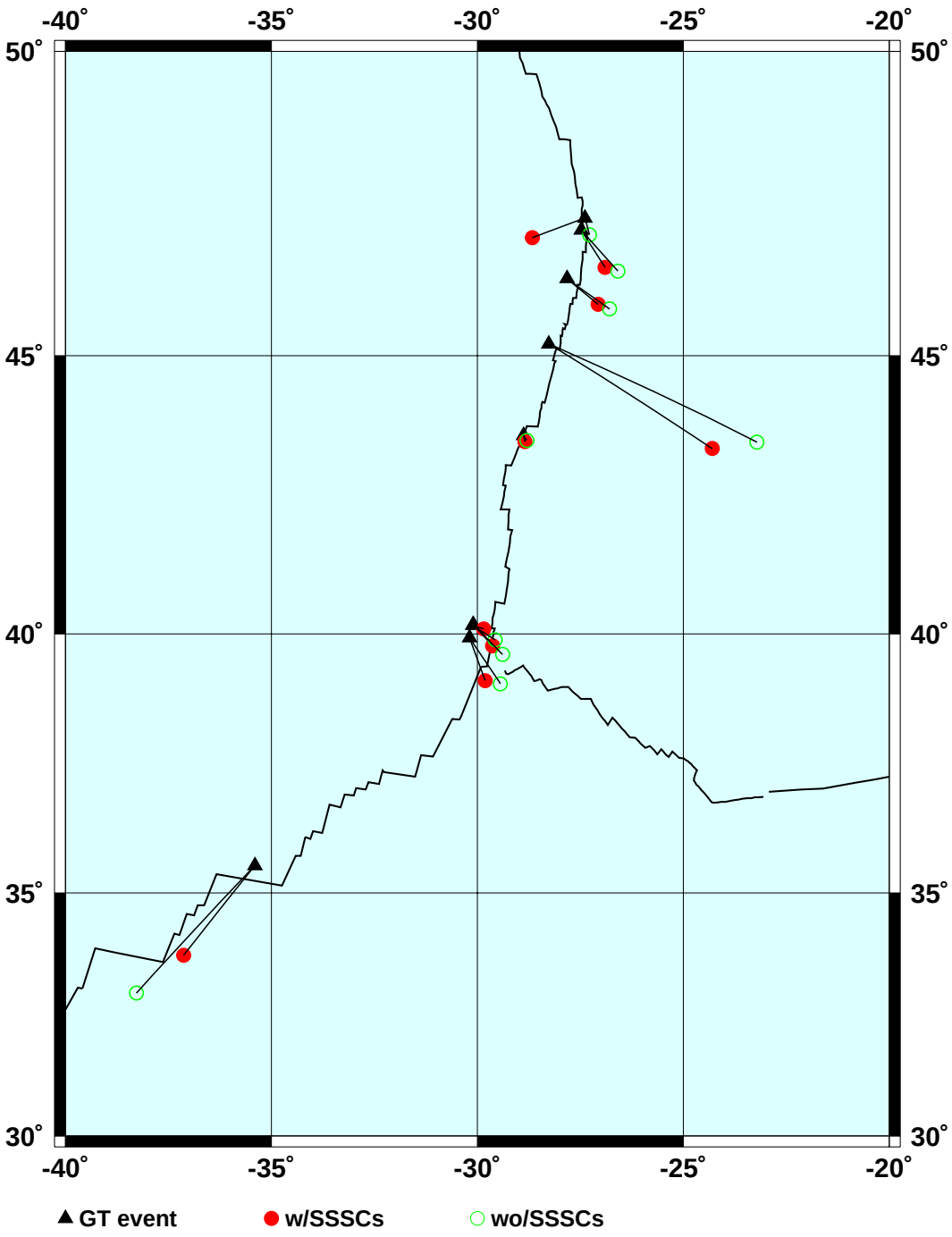

Figure A-22. Location of the MORT GT10 events in North Atlantic when relocated using Pn and Sn phases from all stations within 15°, with and without SSSCs (Appendix 8). Event locations with (solid) and without (open) SSSCs, and GT locations for the 9 GT10 events are all plotted.